PROCEEDINGS
OF THE

NATIONAL ACADEMY OF SCIENCES

OF THE

UNITED STATES OF AMERICA

EXECUTIVE COMMITTEE

OF THE ACADEMY

[illegible] Managing Editor

[illegible] [illegible]

[illegible] Ohio: Mack Printing Company, Easton, Pa.

[illegible] Office: Harvard School of Public Health, Boston, Mass.

[illegible] Office of the Academy: Washington, D. C.

[illegible] at the post-office at [illegible], under the Act of [illegible]

[illegible] authorized [illegible] Act of October 3, 1917, Authorized [illegible]

Price per annum [illegible]

OFFICERS OF THE ACADEMY

PRESIDENT
VICE PRESIDENTS
FOREIGN SECRETARY
HOME SECRETARY
TREASURER

EDITORIAL BOARD

The HOME SECRETARY and the FOREIGN SECRETARY
THE CHAIRMAN of the NATIONAL RESEARCH COUNCIL
A. BRIAN
R. E. CLELAND
A. L. DAY
C. A. KOFOID
L. P. KYROPOULOS
L. C. GRAVE
E. W.
E. B.
A. B.
E. D.
W. E.
W. A.

INFORMATION TO SUBSCRIBERS

Checks covering subscriptions to the PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES at the rate of $6.00 per volume should be made payable to the National Academy of Sciences and addressed to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C. Single numbers $0.50.

Past publications of the NATIONAL ACADEMY OF SCIENCES from 1915 to 1930 and of the NATIONAL RESEARCH COUNCIL from 1919 to 1931 are listed in the Index to the First Ten Volumes of the PROCEEDINGS. This Index appeared as Part 2 of Volume 17, No. 1, January, 1931, and is for sale separately at the price of $1.00. Orders should be sent to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

Past volumes of the PROCEEDINGS are available at five dollars per volume unless the sale of the volume was limited, as in the case of Volumes I to XXX. Single numbers may be obtained at fifty cents each, where the sale of such numbers would not cause a shortage of volumes. Only sets or complete sets of the PROCEEDINGS are available for Volumes 1 to XXIX, price $100.00. Orders should be addressed to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

The following publications are issued by the NATIONAL RESEARCH COUNCIL. Orders and inquiries should be addressed to the Publications Office, National Research Council, 2101 Constitution Ave., Washington 25, D. C.

(A) THE BULLETIN OF THE NATIONAL RESEARCH COUNCIL, published at irregular intervals. Individual numbers may be purchased. A discount of 20% from the list price will be given to libraries of educational institutions. Orders should be issued direct from the Publication Office.

(B) THE REPRINT AND CIRCULAR SERIES OF THE NATIONAL RESEARCH COUNCIL. Individual numbers may be purchased. A discount of 20% from the list price will be given to all who establish standing orders. Orders should be issued direct from the Publication Office.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 36 February 15, 1950 Number 2

A NOTE ON THE MEAN SQUARE VELOCITY IN STELLAR STATISTICS

By Su-Shu Huang

Yerkes Observatory, Williams Bay, Wisconsin, University of Chicago

Communicated by Otto Struve, December 28, 1949

1. Introduction.—In a recent paper Chandrasekhar and Münch[1] have obtained a formula connecting the mean rotational velocity (or its higher moments) of stars with the mean value (or the higher moments) of a directly observable quantity $v \sin i$, where i is the inclination of axis of rotation of the star with the line of sight. The merit of this formula lies in the fact that we can compute a true mean value without any knowledge of the true distribution of the rotational velocity. Here we shall apply this same idea to a different problem in stellar statistics.

Considering stars in a certain region of the sky, we can observe the radial velocity and the two components of tangential velocity;[2] we can therefore obtain from the observational data the mean square velocity components in three mutually perpendicular directions. The question we shall consider is the method by which these means can be converted into mean square velocity components in three principal directions of stellar motion, for example, along the three principal axes of the velocity ellipsoid.[3] Before going into the problem just mentioned, we shall first formulate the basic idea underlying Chandrasekhar and Münch's paper in a more general form.

2. A Statistical Relation.—Let $\phi(y)$ be the observed distribution with respect to a directly observable quantity y, which is a function of two independent physical quantities x and t; thus

$$y = g(x, t). \tag{1}$$

Let the probability that t lies between t and $t + dt$ be $w(t)$. Then the probability distribution, $f(x)$, of the physical quantity x is evidently related to $\phi(y)$ by an integral equation which can be derived in the following manner:

Let y_0 and y_1 denote the least and the greatest possible values of y, respectively. Then

$$g(x, t) = \text{constant} = a \quad (y_0 \leqq a \leqq y_1) \tag{2}$$

represents a one-parameter family of curves in the (x, t) plane. If this family of curves covers continuously a certain domain S in the (x, t) plane without overlapping, then the integral equation relating $\phi(y)$ and $f(x)$ can be verified to be

$$\phi(y) = \int_\alpha^\beta f(x)w(t)\left|\frac{\partial g}{\partial t}\right|^{-1} dx \tag{3}$$

β and α being the boundaries of the domain in the (x, y) plane resulting from the transformation of the (x, t) plane according to (1); hence both β and α may be functions of y. The mapping of the (x, t) plane on the (x, y) plane by (1) is assumed to be one-one and continuous. From (3) we obtain the following relation:

$$\overline{y^n} = \int_{y_0}^{y_1} y^n\phi(y)\, dy = \int_S\int g^n(x, t)f(x)w(t)\, dx\, dt. \tag{4}$$

If the observed quantity y is a function of more than two variables,

$$y = g(x, t_1, \ldots t_n)$$

and if the probability distribution for $x, t_1, \ldots t_n$ are $f(x), w_1(t_1), \ldots w_n(t_n)$ respectively, then

$$\overline{y^n} = \int \ldots \int g^n(x, t_1, \ldots t_n)f(x)w_1(t_1)\ldots w_n(t_n)\, dx\, dt_1 \ldots dt_n. \tag{5}$$

It is worthy of notice that $f(x)$ may also involve the variables t_i.

If the variables in $g(x, t)$ are separable, i.e.,

$$y = g(x, t) = g_1(x)g_2(t)$$

then

$$\overline{y^n} = \int g_1^n(x)f(x)\, dx \int g_2^n(t)w(t)\, dt \tag{6}$$

and we obtain the relation

$$\overline{y^n} = \overline{x^n}\int g_2^n(t)w(t)\, dt \tag{7}$$

for $g_1(x) = x$. An example of (7) is the case considered by Chandrasekhar and Münch; thus, for their problem

$$g_2(t) = w(t) = \sin t.$$

Therefore

$$\overline{y^n} = \overline{x^n}\int_0^{\pi/2} \sin^{n+1} t\, dt.$$

3. Relations between Mean Square Velocities.—Consider a celestial sphere with the observer at the center O. Let OV represent the direction of space velocity of a star at S. Choosing a coordinate system with OX pointing in the direction of star streaming and OZ in the direction of minimum mean speed, we measure the spherical coordinates θ_i and φ_i from the OX-axis and the OXZ-plane, respectively. We let the coordinates of S be (θ_1, φ_1) and of V, (θ_2, φ_2). Furthermore we denote the arc SV by Θ and the angle between the planes OSV and OSX by ϕ.

If $\vec{v}$ denotes the space velocity, the components v_1, v_2 and v_3 of $\vec{v}$ resolved along the radial direction OS, and in the two mutually perpendicular directions in the tangential plane are, respectively,

$$v_1 = v\cos\Theta, \qquad v_2 = v\sin\Theta\cos\phi, \qquad v_3 = v\sin\Theta\sin\phi. \tag{8}$$

According to the general formula (5), the observed mean square velocities (i.e., $\overline{v_1^2}$, $\overline{v_2^2}$ and $\overline{v_3^2}$) are related to the distribution $f(v)$ by the following equations

$$\overline{v^2} = \frac{1}{4\pi}\int\int\int v^2\cos^2\Theta\sin\Theta\, f(v)\, dv\, d\Theta\, d\phi, \tag{9}$$

$$\overline{v_2^2} = \frac{1}{4\pi}\int\int\int v^2\sin^2\Theta\cos^2\phi\, f(v)\, dv\, d\Theta\, d\phi, \tag{10}$$

$$\overline{v_3^2} = \frac{1}{4\pi}\int\int\int v^2\sin^2\Theta\sin^2\phi\, f(v)\, dv\, d\Theta\, d\phi. \tag{11}$$

If the velocity distribution function $f(v)$ were independent of the direction (spherical distribution) equations (9–11) lead to the well-known result

$$\overline{v_1^2} = \overline{v_2^2} = \overline{v_3^2} = \tfrac{1}{3}\overline{v^2}.$$

In fact, as is well known, the velocity distribution does depend upon the direction. Let u_1, u_2 and u_3 be the velocity components along the three coordinate axes, and l, m and n be the direction cosines of OV then

$$v = lu_1 + mu_2 + nu_3$$

and

$$\overline{v^2} = l^2\overline{u_1^2} + m^2\overline{u_2^2} + n^2\overline{u_3^2} \tag{12}$$

if we assume that there is no correlation between two different components so that $\overline{u_iu_j} = 0$ for $i \neq j$. This is the only assumption we need to make and it is of such a general nature that it will include all cases of practical interest. Since

$$l = \cos\theta_2, \qquad m = \sin\theta_2\sin\varphi_2, \qquad n = \sin\theta_2\cos\varphi_2, \tag{13}$$

equations (9–11) can be rewritten in the forms

$$\overline{v_1^2} = A_1\overline{u_1^2} + B_1\overline{u_2^2} + C_1\overline{u_3^2}, \tag{14}$$

$$\overline{v_2^2} = A_2\overline{u_1^2} + B_2\overline{u_2^2} + C_2\overline{u_3^2}, \tag{15}$$

$$\overline{v_3^2} = A_3\overline{u_1^2} + B_3\overline{u_2^2} + C_3\overline{u_3^2}, \tag{16}$$

where A_1, B_1, ..., etc., denote the expressions which can be obtained by inserting (12) in (9–11). Thus,

$$4\pi A_1 = \int\int \cos^2\Theta \sin\Theta \cos^2\theta_2 \, d\Theta \, d\phi, \text{ etc.},$$

where the integration is effected over the entire spherical surface. The various integrations can be carried out if use is made of the relations

$$\sin\theta_2 \cos\varphi_2 = (\cos\Theta \sin\theta_1 - \sin\Theta \cos\theta_1 \cos\phi)\cos\varphi_1 - \sin\Theta \sin\phi \sin\varphi_1 \tag{17}$$

and

$$\sin\theta_2 \sin\varphi_2 = (\cos\Theta \sin\theta_1 - \sin\Theta \cos\theta_1 \cos\phi)\sin\varphi_1 + \sin\Theta \sin\phi \cos\varphi_1 \tag{18}$$

which readily follow from the standard formulas of spherical trigonometry.

In this manner we obtain

$$\left.\begin{array}{lll} 15A_1 = 1 + 2\cos^2\theta_1, & 15A_2 = 1 + 2\sin^2\theta_1, & 15A_3 = 1 \\ 15B_1 = 1 + 2\sin^2\theta_1 \sin^2\varphi_1, & 15B_2 = 1 + 2\cos^2\theta_1 \sin^2\varphi_1, & \\ & 15B_3 = 1 + 2\cos^2\varphi_1 & \\ 15C_1 = 1 + 2\sin^2\theta_1 \cos^2\varphi_1, & 15C_2 = 1 + 2\cos^2\theta_1 \cos^2\varphi_1, & \\ & 15C_3 = 1 + 2\sin^2\varphi_1. & \end{array}\right\} \tag{19}$$

Thus equations (14–16) together with (19) determine $\overline{u_i^2}$ completely from the observed quantities $\overline{v_i^2}$, if we observe only in the immediate neighborhood of a definite point in the sky.

It is also of interest to notice that

$$\overline{v_1^2} + \overline{v_2^2} + \overline{v_3^2} = {}^1/_3(\overline{u_1^2} + \overline{u_2^2} + \overline{u_3^2}), \tag{20}$$

a property independent of (θ_1, φ_1).

If furthermore we observe stars uniformly all over the sky (as for example a definite number of stars per unit area), the mean squares of the observed data will be averages with respect to θ_1 and φ_1 as well as with respect to Θ and ϕ. Multiplying equations (14–16) by $\sin\theta_1 \, d\theta_1 \, d\varphi_1/(4\pi)$ and integrating over the entire celestial sphere, we finally get

$$45\overline{v_1^2} = 5\overline{u_1^2} + 5\overline{u_2^2} + 5\overline{u_3^2}, \tag{21}$$

$$45\overline{v_2^2} = 7\overline{u_1^2} + 4\overline{u_2^2} + 4\overline{u_3^2}, \tag{22}$$

$$45\overline{v_3^2} = 3\overline{u_1^2} + 6\overline{u_2^2} + 6\overline{u_3^2}. \tag{23}$$

As $\overline{u_1^2} \geqq \overline{u_2^2} \geqq \overline{u_3^2}$, it follows from these equations that $\overline{v_2^2} \geqq \overline{v_1^2} \geqq \overline{v_3^2}$; this inequality must be clearly valid since v_2 is always in the direction of maximum mean motion.

Similar results can be obtained for the components of the tangential velocity v_2' and v_3' in the direction of minimum mean motion and in a direction perpendicular to it. For this purpose we choose the directions of minimum mean motion as the X-axis from which θ_1 and θ_2 are measured. Evidently v_1' does not change its meaning but still represents the radial velocity. The final results can be written down by a cyclic change of the indices. Thus for a definite region of the sky we have

$$\overline{v_i'^2} = A_i\overline{u_3^2} + B_i\overline{u_1^2} + C_i\overline{u_2^2} \quad (i = 1, 2, 3), \tag{24}$$

and the equations similar to (21–23) will be

$$45\overline{v_1'^2} = 5\overline{u_3^2} + 5\overline{u_1^2} + 5\overline{u_2^2}, \tag{25}$$

$$45\overline{v_2'^2} = 7\overline{u_3^2} + 4\overline{u_1^2} + 4\overline{u_2^2}, \tag{26}$$

$$45\overline{v_3'^2} = 3\overline{u_3^2} + 6\overline{u_1^2} + 6\overline{u_2^2}. \tag{27}$$

Hence in this case we have

$$\overline{v_3'^2} \geqq \overline{v_1'^2} \geqq \overline{v_2'^2}.$$

We can, of course, also express our results by choosing the direction of u_2 as the X-axis; the final equations will be obtained by a further cyclic change of indices in u_i.

It is of course clear that analogous relations between $\overline{v_i^{2n}}$ and $\overline{u_i^{2n}}$ can also be derived.

4. On a Possible Application of the Formulas of the Previous Section.— If the direction of star streaming and the direction of minimum mean motion are assumed to be known, we may compute the axes of the velocity ellipsoid by means of the relations derived. Among the three components of velocity v_1, v_2 and v_3 of the stars, the radial velocity v_1 is the one which is most accurately determined. Therefore in practical applications it will be advisable to utilize only the observational material on the radial velocities. This can be achieved by considering (14) only without using (15, 16).

Dividing the whole sky into small regions of, say 100 square degrees we use for each region the equation

$$15\overline{v_{rad}^2} = (1 + 2\,\overline{\cos^2\theta_1})\overline{u_1^2} + (1 + 2\,\overline{\sin^2\theta_1 \sin^2\varphi_1})\overline{u_2^2} + (1 + 2\,\overline{\sin^2\theta_1 \cos^2\varphi_1})\overline{u_3^2}. \tag{28}$$

Equation (28) is simply (14) averaged over the small region. We shall thus obtain as many equations as the number of regions into which the whole sky has been divided. (In practice we may choose only those regions where the number of stars with measured radial velocities exceeds a certain lower

limit and discard others for statistical reasons.) A least squares solution of these equations will determine the three unknowns $\overline{u_1^2}$, $\overline{u_2^2}$ and $\overline{u_3^2}$. In forming the normal equations the number of stars in each region should enter as the weight factor.

This short note was inspired by Professor S. Chandrasekhar's stimulating colloquium on the paper cited previously; thus I should like to express my sincere thanks to him. It is also my pleasure to put on record that equation (7) was independently noticed by both Dr. A. Brown and D. E. Osterbrock. To both of them I am also indebted for some valuable discussions.

[1] Chandrasekhar, S., and Münch, G., to appear in January, 1950, issue of *Astrophys. J.*

[2] Due to the statistical nature of the present problem, it suffices to use the spectroscopic parallax in order to obtain the tangential velocity.

[3] For the terminology used in the present communication see Chandrasekhar, *Principles of Stellar Dynamics*, University of Chicago Press, 1942.

---

*RADIAL OSCILLATIONS OF COMPRESSIBLE GAS SPHERES**

BY ZDENĚK KOPAL

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND HARVARD COLLEGE OBSERVATORY

Communicated by Harlow Shapley, November 25, 1949

The problem of radial oscillations of compressible gas spheres in hydrostatic equilibrium has so far been solved analytically for four different model configurations, three of which were discovered by Sterne,[1] while the fourth one was recently added by Prasad.[2] In each case, the respective configuration proved to be capable of oscillating in a discrete set of frequencies dependent on the mean density and the ratio of specific heats of the material constituting the configuration. The writer has, however, recently pointed out[3] that the assumptions made by Sterne and Prasad concerning the structure of their models were, in each case, such as to reduce the differential equations to the hypergeometric form; the discrete character of the frequency spectra followed as a consequence of the fact that the respective hypergeometric series of unit radius were found to be divergent and had to be reduced to polynomials in order to maintain no variation in pressure over the free surface. Although the four models considered by Sterne and Prasad are characterized by outwardly very different distribution of density in their interiors, mathematically they all belong to the same type. This similarity, in turn, prompts us to inquire as to the possible existence of other models, of different constitution, whose eigen-amplitudes of radial oscillations may also be expressible in terms of hypergeometric series. Are the four known models the only ones possessing this property, or are there others of this class which have so far escaped discovery?

The object of this investigation will be to supply an exhaustive answer to this question and to provide a complete enumeration of the respective configurations. It will be shown that, for finite values of central condensation (i.e., of the ratio of the mass stored at the center to that of the whole configuration), no models other than those investigated by Sterne and Prasad exist for which the amplitudes of small radial oscillations are of the hypergeometric type. If, however, the degree of central condensation is allowed to increase without limit, a whole new family of such models is found to exist which is characterized by a partly discrete and partly continuous spectrum of the frequencies of free oscillation. This latter branch of the family represents the first known instance of compressible gas spheres, in hydrostatic equilibrium, which can perform small radial oscillations in *any* frequency.

1. Equations of the Problem.—As is well known[4] the differential equation governing the variation of the amplitude $\xi(x)$ of small adiabatic radial oscillations in the interior of a gas configuration in hydrostatic equilibrium can be reduced to the form

$$\frac{d^2\xi}{dx^2} + \left\{\frac{4}{x} + \frac{d\log P}{dx}\right\}\frac{d\xi}{dx} + \left\{\frac{\bar{\alpha}}{x} - \frac{n^2R}{\bar{\gamma}g}\right\}\frac{d\log P}{dx}\,\xi = 0, \qquad (1)$$

where P denotes the pressure; g, the gravity; n, the frequency of the oscillation; $\bar{\gamma}$, the ratio of specific heats; $\bar{\alpha} = 3 - (4/\bar{\gamma})$; and $x = r/R$, r being the distance of any arbitrary point from the center of a spherically symmetrical configuration of radius R. Let us assume now that, by hypothesis,

$$\xi(x) \propto x^c F(\alpha, \beta, \gamma, ax^b), \qquad (2)$$

where F denotes the ordinary hypergeometric series and a, b, c, α, β, γ are arbitrary constants. The function on the right-hand side of (2)—let us call it y—is known to satisfy the equation

$$\frac{d^2y}{dx^2} + \frac{apx^b + q}{x(ax^b - 1)}\frac{dy}{dx} + \frac{arx^b + s}{x^2(ax^b - 1)}\,y = 0, \qquad (3)$$

where

$$\left.\begin{aligned}
\alpha &= \frac{1}{b}\left\{\frac{p+q}{2} + \sqrt{\left(\frac{q+1}{2}\right)^2 + s} + \sqrt{\left(\frac{p-1}{2}\right)^2 - r}\right\} \\
\beta &= \frac{1}{b}\left\{\frac{p+q}{2} + \sqrt{\left(\frac{q+1}{2}\right)^2 + s} - \sqrt{\left(\frac{p-1}{2}\right)^2 - r}\right\} \\
\gamma &= 1 + \frac{2}{b}\sqrt{\left(\frac{q+1}{2}\right)^2 + s},
\end{aligned}\right\} \qquad (4)$$

while c is the root of the quadratic equation

$$c^2 - (q + 1)c - s = 0. \tag{5}$$

If there is to be no displacement at the center—which is one of the boundary conditions of our problem—only positive roots of this equation are of physical interest.

The series on the right-hand side of (2) is known to converge absolutely and uniformly for $ax^b < 1$ or, should $ax^b = 1$, if $\alpha + \beta < \gamma$. It will be divergent if $ax^b > 1$ or, for $ax^b = 1$, if $\alpha + \beta \geqslant \gamma$. An appeal to equations (4) discloses that, if $ax^b = 1$, our series will be divergent if

$$p + q \geqslant b, \tag{6}$$

and will be convergent if the opposite is true. If, by hypothesis, $y = \xi(x)$, and equations (1) and (3) are, therefore, to be identical, the structure of the oscillating configuration must evidently be such that

$$\frac{ap x^b + q}{x(ax^b - 1)} = \frac{4}{x} + \frac{d \log P}{dx} \tag{7}$$

and

$$\frac{arx^b + s}{x^2(ax^b - 1)} = \left\{\frac{\bar{\alpha}}{x} - \frac{n^2 R}{\bar{\gamma} g}\right\} \frac{d \log P}{dx}, \tag{8}$$

respectively. Equation (7) implies that

$$\frac{d \log P}{dx} = \frac{a(p - 4)x^b + q + 4}{x(ax^b - 1)}, \tag{9}$$

which can be integrated into

$$P = k(1 - ax^b)^{(p + q)/b} x^{-(q + 4)}, \tag{10}$$

where k is a constant. Since the pressure must, by definition, vanish on the surface where $x = 1$, equation (10) makes it evident that this can be true if, and only if, $a = 1$. The first of the six arbitrary constants occurring in equation (3) has thus been specified.

Next let us eliminate the logarithmic derivative of P between (7) and (8) and solve the resulting equation for g: we obtain

$$g = \frac{x(Cx^b + D)}{Ax^b + B}, \tag{11}$$

where we have abbreviated

$$\left.\begin{aligned} A &= \bar{\gamma}[\bar{\alpha}(p - 4) - r] \\ B &= \bar{\gamma}[\bar{\alpha}(q + 4) - s] \\ C &= n^2 R(p - 4) \\ D &= n^2 R(q + 4). \end{aligned}\right\} \tag{12}$$

Now, as is well known,

$$g = G\frac{m(x)}{R^2x^2}, \tag{13}$$

where G denotes the constant of gravitation and $m(x)$, the mass of our configuration interior to x. Provided that the density is a continuous function of x,

$$m(x) = 4\pi GR^3\int_0^x \rho x^2\,dx. \tag{14}$$

Inserting (11) in (14) and differentiating with respect to x we find that the density inside of our configuration should vary as

$$\rho(x) = \frac{3ACx^{2b} + [(3-b)AD + (3+b)BC]x^b + 3BD}{4\pi GR(Ax^b + B)^3}. \tag{15}$$

All constants in this equation are so far arbitrary—save for the obvious requirement that they be such as to render $\rho(x)$ positive throughout the interior. The central density of such a configuration is given by

$$\rho_c = \frac{3D}{4\pi GRB}, \tag{16}$$

while its mean density becomes

$$\bar{\rho} = \frac{3}{4\pi GR}\frac{C+D}{A+B}. \tag{17}$$

Lastly, let us ensure that the configuration, in which the pressure, density, and gravity are governed by equations (10), (11) and (15), can be in hydrostatic equilibrium. As is well known, this will be the case provided that P, ρ and g are related by

$$\frac{dP}{dr} = \frac{1}{R}\frac{dP}{dx} = -g\rho. \tag{18}$$

If we differentiate (10) and insert in (18) together with (11) and (15), the equation of hydrostatic equilibrium will take the explicit form

$$\frac{4\pi kG}{3n^2R}\frac{B^2}{D} = \frac{1 + \left[\left(1-\frac{b}{3}\right)\frac{A}{B} + \left(1+\frac{b}{3}\right)\frac{C}{D}\right]x^b + \frac{AC}{BD}x^{2b}}{x^{-(q+b)}(1-x^b)^{[(p+q)/b]-1}\left(1+\frac{A}{B}x^b\right)^3} \tag{19}$$

and must hold good throughout the interior (except, possibly, at the center). This latter equation supplies the remaining set of conditions

which the arbitrary constants in equation (3) must fulfill in order to provide a physically admissible solution. In what follows we shall, therefore, set out to ascertain all possible combinations of the values of b, p, q, r and s which conform to these conditions.

2. *Equilibrium Configurations.*—An inspection of equation (19) makes it evident that, if all four constants A, B, C, D as well as b are to be real and different from zero, this equation cannot be satisfied for any value of x; for if all four were non-vanishing, equation (19) would require that

$$\left.\begin{aligned} (D/B)^2 &= (4/3)\pi G k(q+4) \\ q+6 &= 0 \end{aligned}\right\} \tag{20}$$

be satisfied simultaneously, and this would render at least one of our constants imaginary. A closer examination discloses, moreover, that the same situation prevails if *one* of the four constants A, B, C, D is set equal to zero—except when $B = 0$ and $b = 3$, in which case Prasad's model follows. It is not until *pairs* of these constants are permitted to vanish simultaneously that we obtain all other configurations which are consistent with our initial assumption (2).

If $A = C = 0$ or $A = D = 0$, one of the two remaining constants still turns out to be imaginary. If $A = B = 0$, the corresponding configuration would be one of infinite mass, density and gravity; if $C = D = 0$, the mass of our configuration would be zero. The case of $B = D = 0$ leads to a homogeneous configuration (Sterne's "Model 1"), while if $B = C = 0$ two different types of configurations are possible. When $b = 2$, a model is obtained in which the density varies continuously as x^{-2}; the central density is infinite, but the mass of the whole configuration remains finite (Sterne's "Model 2"). If, however, $b = 3$, the density is discontinuous at $x = 0$ and such that the whole mass of the configuration will be stored at the center; the weight of the surrounding envelope being infinitesimal the gravity falls off with the inverse square of the distance from the center. A whole family of such models, characterized by $-4 < q < 0$, is found to be consistent with the assumption of hydrostatic equilibrium, but only one particular member of it (corresponding to $q = -1$) has so far been noticed (Sterne's "Model 3"). In what follows, all these models will be discussed in turn.

3. *Homogeneous Models.*—If, in equations (12), the values of our arbitrary constants q and s are chosen so as to render $B = D = 0$, the density of the respective configuration as defined by equation (15) turns out to be constant and equal to

$$\rho = \frac{3C}{4\pi GRA} = \frac{3n^2(p-4)}{4\pi G\bar{\gamma}[\bar{\alpha}(p-4)-r]}, \tag{21}$$

and an appeal to equations (16) and (17) discloses that

$$\rho_c = \rho(x) = \bar{\rho}. \tag{22}$$

Now the condition $D = 0$ implies, by (12), that $q = -4$ and, consequently, $B = 0$ implies that $s = 0$. Moreover, equation (19) safeguarding the existence of hydrostatic equilibrium requires that $k = (2/3)\pi G\bar{\rho}^2R^2$, $b = 2$ and $p = 6$. Inserting this latter in (21) we find ultimately that

$$r = \frac{3n^2}{2\pi G\bar{\rho}\hat{\gamma}} - 2. \tag{23}$$

Equations (10), (11) and (15) describing the model under investigation become

$$\left.\begin{aligned} P &= \frac{2}{3}G\bar{\rho}^2R^2(1 - x^2) \\ \rho &= \bar{\rho} \\ g &= \frac{4}{3}\pi G\bar{\rho}Rx. \end{aligned}\right\} \tag{24}$$

If this model were disturbed slightly from its state of equilibrium in such a way that a small purely radial motion results, its amplitude $\xi(x)$ would be of the form (2) where the constants α, β and γ are obtained from equations (4) by inserting in them the foregoing values of p, q, r and s; while equation (5) yields, for $q = -4$, $c = 0$ or -3. A requirement that there be no displacement at the center rules out the negative root; hence, $c = 0$. The outer boundary condition requiring that there be no variation of pressure on the surface can be met only if $\xi(1)$ is a finite quantity. Now the values of b, p and q in the present case turned out to be such that $p + q = b$, in which case the criterion (6) discloses that the hypergeometric series on the right-hand side will diverge for $x = 1$. If $\xi(1)$ is to be a finite quantity, we must reduce the respective infinite series to a polynomial by making α or β equal to zero or some negative integer by setting

$$r = -2j(2j + 5), \qquad j = 0, 1, 2, \ldots \tag{25}$$

which, combined with (21), specifies the characteristic frequencies n in which our configuration can oscillate freely in the jth mode. The corresponding amplitudes of oscillation then take the explicit forms

$$\xi_j(x) \propto G_j\left(\frac{5}{2}, \frac{5}{2}, x^2\right), \tag{26}$$

where $G_j(p, q, x) \equiv F(p + j, -j, q, x)$ denotes the respective Jacobi polynomial.[5]

The conditions $B = D = 0$ are not the only ones which lead to equation (21) for, as the reader can easily verify, a combination $B = 0$ and $b = 3$

will reduce (15) to exactly the same result. We thus obtain another homogeneous configuration which is, however, different from the preceding one; for the equation (19) of hydrostatic equilbrium now calls for $p = 6$ but $q = -3$, in virtue of which the condition $B = 0$ implies that $s = \bar{\alpha}$, and

$$k = \frac{3n^4R^2}{2\pi GA^2}, \tag{27}$$

where

$$A = \gamma(2\bar{\alpha} - r). \tag{28}$$

On the other hand, equation (17) defining the mean density of our configuration takes the explicit form

$$\bar{\rho} = \frac{9n^2}{4\pi GA}. \tag{29}$$

Eliminating A between (27) and (28) by means of (29) we obtain

$$r = 2\bar{\alpha} - \frac{9n^2}{4\pi G\bar{\gamma}\rho}, \tag{30}$$

and

$$k = \frac{8}{27}\pi GR^2\bar{\rho}^2. \tag{31}$$

A comparison of equations (21) and (29) discloses that

$$\rho(x) = \frac{2}{3}\bar{\rho}, \tag{32}$$

while equation (10) assumes the explicit form

$$P = \frac{8}{37}\pi G\bar{\rho}^2R^2\frac{1 - x^2}{x}. \tag{33}$$

Both the preceding equations are different from (24) and the reason is not too hard to find. Unlike the preceding model, the validity of equation (21) must be limited to $0 < x < 1$, with the center $x = 0$ being a singular point. Our configuration must consist of a central mass-point, of mass m_0, surrounded by an envelope of constant density and mass m_e. The ratio of these two masses is already fixed by equation (32); for, by definition, we have

$$\frac{\rho(x)}{\bar{\rho}} = \frac{m_e}{m_0 + m_e} \tag{34}$$

which, combined with (32), yields

$$m_e = 2m_0. \tag{35}$$

Hence, the homogeneous envelope should contain two-thirds of the total mass of our configuration. The total gravity prevailing in our configuration will eventually be obtained from equation (11), the right-hand side of which is to be increased by the term Gm_0/R^2x^2 arising from the central mass; doing so we obtain

$$g(x) = \frac{m}{3R^2}\left\{2x + \frac{1}{x^2}\right\}, \tag{36}$$

where $m = m_0 + m_e$ denotes the total mass of our configuration.

If this model were disturbed slightly from a state of equilibrium in such a way that a small purely radial motion results, its amplitude would be of the form (2) where the constants α, β and γ are obtained by inserting in equations (4) the foregoing values of p, q, r and s, while the positive root of equation (5) is

$$c = \sqrt{1 + \tilde{\alpha}} - 1. \tag{37}$$

As in the preceding case, we again have $\alpha + \beta = \gamma$ and the hypergeometric series $F(\alpha, \beta, \gamma, 1)$ again diverges. If ξ is to remain finite on the surface, therefore, the series must be terminated by the condition:

$$r = -(3j + c)(3j + c + 5), \tag{38}$$

where c is given by the preceding equation (37) and j is zero or a positive integer. Equation (38) combined with (30) specifies the characteristic frequencies n in which our configuration can freely oscillate in jth mode; the corresponding amplitudes take the explicit form

$$\xi_j(x) \propto x^c G_j\left(\frac{5}{3} + \frac{2}{3}c, \frac{5}{3} + \frac{2}{3}c, x^2\right) \tag{39}$$

where G_j denotes, as before, the respective Jacobi polynomial.

4. Heterogeneous Models.—The foregoing two homogeneous models characterized by $m_0/m_e = 0$ and 0.5, respectively, are the only types of configurations, consistent with our initial assumption (2), in which $\rho(x) =$ constant for $0 < x < 1$. In the remaining cases which yield a closed solution, the density will vary throughout the interior and our aim will be to choose this variation so as to make (2) a solution of our fundamental equation. Two kinds of such models will be found to exist: one exhibiting a finite degree of central condensation, the other consisting of a central point-mass surrounded by an envelope of infinitesimal weight and certain specific structure. This latter family of models will be found to possess

properties which are particularly interesting; but for the sake of completeness we shall take up the former and simpler model first.

This configuration is obtained by setting $B = C = 0$ or, which is the same, $p = 4$ and $s = 2\bar{a}$. The equation (19) of hydrostatic equilibrium will then be satisfied if $b = 2$, $q = -2$ and

$$k = \frac{D^2}{8\pi GA^2} = \frac{2}{9}\pi GR^2\bar{\rho}^2 \tag{40}$$

by (17). Moreover, equations (12) then yield $A = -r$ and $D = 2n^2R$ which, combined with (17) gives

$$r = \frac{3n^2}{2\pi G\bar{\rho}}. \tag{41}$$

Our configuration will be therefore characterized by

$$\left.\begin{aligned} P &= \frac{2}{9}\pi GR^2\bar{\rho}^2\,\frac{1-x^2}{x^2}, \\ \rho &= \frac{\bar{\rho}}{3x^3}, \\ g &= \frac{4\pi GR\bar{\rho}}{3x}. \end{aligned}\right\} \tag{42}$$

Inside our model the density is found to vary as the inverse square of the distance from the center and to increase beyond any limit as $x \to 0$. This singularity is, however, integrable; for the mass of the configuration is clearly finite.

The constants α, β, γ characterizing the hypergeometric series on the right-hand side of (2) are obtained if we insert the above values of p, q, r and s in equations (4), while the positive root of (5) for $q = -2$ becomes

$$c = \frac{1}{2}\{\sqrt{1 + 8\bar{a}} - 1\}. \tag{43}$$

The criterion (6) discloses that, for $p = 4$, $q = -2$, and $b = 2$, our hypergeometric series is again divergent where $x = 1$; hence, if the amplitude of the oscillations is to remain finite on the surface the series must be terminated by setting

$$r = (2j + c)(2j + c + 3), \tag{44}$$

where j is zero or a positive integer. The eigen-amplitudes (2) reduce then to Jacobi polynomials of the form

$$\xi_j(x) \propto x^c G_j\left(\frac{3}{2} + c, \frac{3}{2} + c, x^2\right), \tag{45}$$

while the characteristic frequencies of oscillation n are defined by the preceding equation (44).

In all models analyzed so far, the mass inside our configuration turned out to be a continuous and increasing function of x, with all layers contributing significantly to the effective gravity. In the final part of our discussion, a family of models will be considered in which the overwhelming part of the mass of the whole configuration is stored at its center so that, in effect, $m(x) = m(R)$ and the gravity inside of such a configuration varies, by definition, as

$$g = \frac{Gm(R)}{R^2x^2}. \tag{46}$$

A comparison of this result with equation (11) discloses that if (46) is to be true, it is necessary that $B = C = 0$, i.e., $p = 4$, $s = \bar{\alpha}(q + 4)$, $b = 3$ and

$$\frac{D}{A} = G\frac{m}{R^2} = \frac{4}{3}\pi G\bar{\rho}R. \tag{47}$$

Under these conditions we again have $A = -r$ which, combined with the preceding equation (47), yields

$$r = -\frac{3(q + 4)n^2}{4\pi G\bar{\rho}}. \tag{48}$$

Moreover, by (10)

$$P = \lim_{k \to 0} k(1 - x^3)^{(q+4)/3}x^{-(q+4)}; \qquad x > 0. \tag{49}$$

Consistent with the assumption of the mass of the envelope surrounding the point-core to be infinitesimal, equation (15) yields indeed $\rho = 0$ (for $x \neq 0$). In order to ascertain the *variation* of this infinitesimal density within the envelope we must, therefore, fall back on the equation (18) of hydrostatic equilibrium which asserts that

$$\rho = -\frac{1}{g}\frac{dP}{dr} = \lim_{k \to 0} \frac{3(q + 4)k}{4\pi GR^2\bar{\rho}} \frac{(1 - x^3)^{[(q+4)/3]-1}}{x^{q+3}}; \qquad x > 0. \tag{50}$$

Equations (49) and (50) render indeed the pressure and density infinitesimal everywhere except at the center where they must be, by definition, infinite.

In all preceding equations the value of q has so far been wholly arbitrary. The only restriction which we shall have to impose upon q will be a requirement that the total mass of the envelope be also a quantity of the order of k—i.e., that the integral on the right-hand side of (14) be convergent for $x = 1$. Explicitly,

$$m(R) = \frac{3(q+4)Rk}{G\beta}\int_0^1 (1-x^3)^{[(4+q)/3]-1}x^{-(q+1)}\,dx$$

$$= \frac{(q+4)Rk}{G\beta}\,B\left(\frac{q+4}{3}, -\frac{q}{3}\right), \tag{51}$$

where B denotes the complete beta-function. The latter is known to be convergent if both its arguments are positive; and they will be such if

$$0 > q > -4. \tag{52}$$

The constants α, β and γ of the hypergeometric series in (2) expressing the amplitudes of our oscillation problem are again obtained when we insert the above values of p, q, r and s in equations (4). Equation (5) takes now the explicit form

$$c^2 - (q+1)c - \bar{\alpha}(q+4) = 0. \tag{53}$$

Its coefficients make it evident that, as long as $q > -4$, one of its two roots will be positive and equal to

$$c = \frac{q+1}{2} + \sqrt{\frac{(q+1)^2}{2} + \bar{\alpha}(q+4)}. \tag{54}$$

The whole family of models corresponding to the range of q bounded by the inequality (53) can, for the purpose of description, be divided into three parts:

$$(a)\quad 0 > q > -1\text{:}\quad \frac{q+4}{3} > 1;\quad \rho(1)/k = 0.$$

$$(b)\quad q = -1\text{:}\quad \frac{q+4}{3} = 1;\quad \rho(1)/k = \text{constant}.$$

$$(c)\quad -1 > q > 4\text{:}\quad \frac{q+4}{3} < 1;\quad \rho(1)/k = \infty.$$

The criterion (6) discloses that if $0 > q > -1$ [which includes the models (a) and (b)] our hypergeometric series diverges when $x = 1$ and must, therefore, be terminated by putting

$$r = -\frac{3(q+4)n^2}{4\pi G\bar{\gamma}\beta} = -(3j+c)(3j+c+3), \tag{55}$$

where c is given by (54) and j stands for zero or a positive integer. The eigen-amplitudes of our problem then take the form

$$\xi_i(x) \propto x^i G_i\left(1 + \frac{2}{3}c, 1 + \frac{2}{3}c, x^2\right), \tag{56}$$

while the eigen-frequencies n are given by the foregoing equation (55). If $0 > q > -1$, equation (50) shows that the density becomes zero when $x = 1$; the reader may note that, of all models discussed so far, this is the only one in which the density as well as pressure vanish on the surface of our configuration. In the limiting case of $q = -1$ the density on the surface becomes arbitrary; this case was previously discovered by Sterne[1] (as his "Model 3").

If, however, q becomes less than -1, the criterion (6) discloses that the hypergeometric series will converge for an *arbitrary* value of r. The right-hand side of equation (2) automatically satisfies, therefore, our outer boundary condition requiring the finiteness of $\xi(x)$ on the surface without imposing any restriction on n—which means that *the corresponding configuration can perform free radial oscillations in any frequency.* This is the first instance of a "continuous spectrum" in the distribution of eigen-frequencies encountered in connection with any gaseous configuration in hydrostatic equilibrium known so far, and the properties of such configurations deserve evidently a close attention. Equation (50) discloses that their density $\rho(x)$, infinite at the center, becomes infinitesimal for $0 < x < 1$, but infinite again for $x = 1$, thus giving rise to an infinitesimally thin surface shell. The mass enclosed in this shell remains, however, a quantity of the order of k and does not, therefore, contribute appreciably to the effective gravity; our basic assumption which led us to (46) has not been violated. As long as $-1 > q > -3$, the density between the center and the surface shell diminishes at first with increasing x until the distance has been reached at which $2x^2 = q + 3$. At this point the density gradient reverses its sign and the density increases hereafter again until the surface is reached. Ultimately, if $-3 \geqslant q > -4$, the density gradient will be positive for $x > 0$ throughout the configuration.

5. Concluding Remarks.—Whether or not the physical situation analyzed in preceding paragraphs finds an application to actual astronomical bodies is an intriguing object for speculation. Our results may possibly have some bearing on the behavior of early-type giant or supergiant stars, in which the "surface shell" may be constituted by the condensing material driven out by the radiation pressure. Several theories of the structure of the giants along such lines were advanced in recent decades, the latest and most satisfactory one being due to Menzel.[6] If our configurations have anything to do with such stellar models, our results would imply that the models are hydrostatically stable under any small and spherically symmetrical disturbance, and would respond to it by setting up pulsations in an *arbitrary* period. For such stars, the product $P\sqrt{\bar{\rho}}$

(P being the pulsation period) could, therefore, assume *any* value. Moreover, if the disturbing force were periodic, a resonance of the free and forced periods would be immediately established and the amplitude of the oscillation greatly enhanced. Second-order effects, ignored in the present investigation, would thus probably soon become appreciable and could operate to destroy the resonance. Should they fail, however, to do so, the only way in which a star built up according to our model could escape from its predicament would be through a gradual change of its whole internal structure which would convert it into another model characterized by a discrete spectrum of eigen-frequencies, or the old model might drift out of the region of continuous frequency spectrum. Barring such an escape in time, however, the consequences of any periodic disturbance of the configurations built up according to our model might become conspicuous, or even cataclysmic—whatever the period of the disturbing force. The extent to which any such phenomena may actually occur in Nature must, however, be left for future investigations to decide.

* Work completed under Contract N5 ori—07843 with the Office of Naval Research.

[1] M. N., **97**, 582 (1937).

[2] *Ibid.*, **108**, 414 (1948).

[3] *Ap. J.*, **111**, No. 1 (January, 1950); in press.

[4] Cf., for instance, ref. 1. The equation was first derived by Eddington, *M. N.*, **79**, 2 (1918).

[5] In the notations of Courant-Hilbert, *Methoden der Matematischen Physik*, Berlin, 1931, vol. I, p. 77.

[6] Zeeman Congress, Amsterdam (1946); *Physica*, **12**, 768 (1946).

THE SYNTHESIS OF RHODOPSIN FROM RETINENE$_1$

BY GEORGE WALD AND PAUL K. BROWN*

BIOLOGICAL LABORATORIES OF HARVARD UNIVERSITY, CAMBRIDGE

Communicated by A. S. Romer, December 4, 1949

Rhodopsin, the light-sensitive pigment of rod vision, is bleached in the retina by light and is continuously resynthesized by ordinary "dark" reactions. Its restoration permits vision to continue in the light, and is the source of visual dark adaptation.

Some seventy years ago Kühne observed that rhodopsin is synthesized in two ways—a rapid regeneration from yellow products of bleaching, persisting in the isolated retina and even to some degree in solution; and a much slower synthesis from colorless precursors, found ordinarily only in the intact eye, and requiring—so Kühne believed—the cooperation of substances from the pigment epithelium.[1] These processes have since been

identified, respectively, as the synthesis of rhodopsin from retinene$_1$ and from vitamin A_1.[2]

Kühne's observation that rhodopsin is restored to some extent after bleaching in solution was confirmed by Hecht, *et al.*,[3] and Chase and Smith.[4] We can conclude from the way these contemporary experiments were performed that regeneration proceeded from the usual end-products of bleaching rhodopsin in solution, retinene$_1$ and protein. It was reported to follow the course of a first-order reaction, and to be optimal at pH 6.7. The largest regeneration recorded was about 15 per cent.

The bleaching of rhodopsin in solution is a complex process. An initial light reaction forms an orange-red product (lumi-rhodopsin) which is transformed by ordinary thermal reactions to a final mixture of retinene$_1$ and protein.[5] The conditions for the further reduction of retinene$_1$ to vitamin A_1—the presence of retinene reductase and reduced cozymase—are not realized in the solutions here considered.[6] As retinene$_1$ forms it is redistributed. In part it remains attached to the groups on rhodopsin-protein on which it originates, in part it couples with other groups on this and other molecules.[6, 7] This wandering of retinene$_1$ away from its original sites of attachment to rhodopsin-protein is an important factor limiting regeneration.

We have indeed good evidence that all stages in the bleaching of rhodopsin to retinene$_1$ are reversible; but that, the further bleaching has advanced, and the more time has been given for retinene$_1$ to leave rhodopsin-protein, the less regeneration occurs. Thus if the bleaching process is blocked in its initial stages (lumi-rhodopsin, meta-rhodopsin), as in extreme cold or dry gelatin films, regenerations of about 50 per cent are observed;[5] but after about an hour of illumination at room temperature in solution, only a few per cent of rhodopsin is regenerated.

These considerations led us to examine the effect of flooding rhodopsin solutions with synthetic retinene$_1$. The retinene$_1$ was prepared by the chromatographic oxidation of crystalline vitamin A_1 on solid manganese dioxide.[8] It was added in 2 per cent aqueous digitonin, the detergent used to extract rhodopsin from the retina. When rhodopsin is bleached in the presence of a high concentration of retinene$_1$, regenerations of about 70 per cent are regularly observed.

Such an experiment is shown in figure 1. A preparation of cattle rhodopsin was divided into three portions. To one a large excess of retinene$_1$ was added, and this and a second portion were exposed for one minute to intense white light. The extinction at 500 mμ—the absorption maximum of rhodopsin—was thereafter measured in darkness. In the solution flooded with retinene$_1$ the extinction rose rapidly as rhodopsin regenerated; in the untreated solution, after a small preliminary rise, the extinction fell due to secondary "dark" components of bleaching. When all changes were

nearly complete, hydroxylamine was added to these solutions and to the third portion of the original rhodopsin, to block any further regeneration (see below).[9] Absorption spectra were measured in the dark, then again after bleaching the three solutions in the light. The differences in absorption represent the rhodopsin present originally ($A - B$), and that regen-

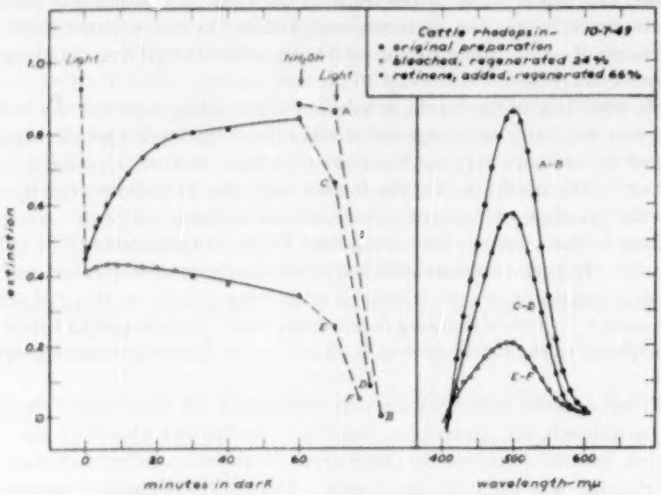


FIGURE 1

Regeneration of rhodopsin in the presence of added retinene$_1$. 25°C., pH 6.2. A highly purified solution of cattle rhodopsin was divided into 3 portions. To one, a high concentration (ca. 30 μg. per ml.) of synthetic retinene$_1$ was added (solid circles); and this and a second untreated portion (open circles) were bleached in white light. Both were then left dark for 60 minutes. Measurements at the left show extinctions at 500 mμ, the wave-length of maximum absorption of rhodopsin. The solution to which retinene$_1$ was added shows a large rise in extinction at this wave-length, the other a net fall. Finally hydroxylamine (0.25 M) was added to these solutions and to the third portion of the original rhodopsin, to block further regeneration. Absorption spectra were measured in the dark; then all three solutions were bleached, and their absorption spectra remeasured. The differences in absorption, at 500 mμ on the left, and throughout the spectrum on the right, represent the rhodopsin present originally ($A - B$), and regenerated in the untreated solution ($E - F$) and in the solution to which retinene$_1$ had been added ($C - D$).

erated in the untreated solution ($E - F$) and in the solution to which retinene$_1$ had been added ($C - D$).

The retinene$_1$ added in such an experiment might act in a number of ways. It could tend to reverse the bleaching process in its intermediate

stages; or to retard the migration of retinene$_1$ away from rhodopsin-protein by occupying all other groups in the solution to which retinene$_1$ could attach. It probably does these things; but it also enters directly into the synthesis of new rhodopsin.

This was demonstrated in experiments in which colorless rhodopsin-protein, free of all native retinene$_1$, was mixed with synthetic retinene$_1$ in high concentration. This mixture, placed in the dark, yields a large synthesis of rhodopsin.

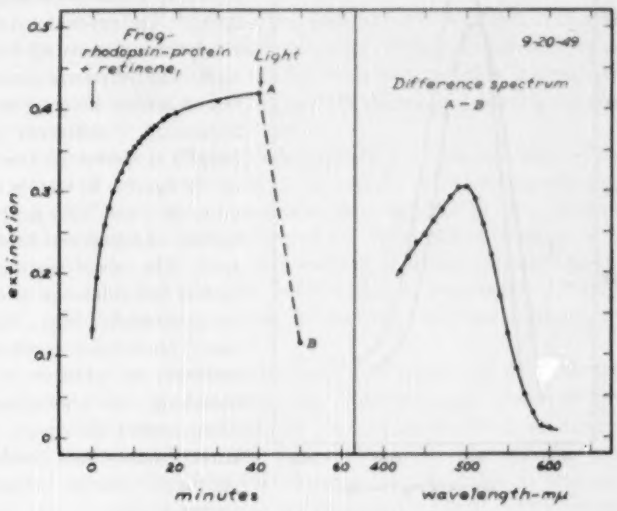


FIGURE 2

Synthesis of rhodopsin from rhodopsin-protein and synthetic retinene$_1$. A colorless solution of rhodopsin-protein extracted from frog retinas which had bleached for 1½ hours in daylight was mixed with synthetic retinene$_1$ to a final concentration of about 10 μg. per ml. The measurements at the left show the rise in extinction in darkness at 500 mμ as rhodopsin is synthesized (25°C., pH 6.3). At *A* the product was exposed to daylight for 20 minutes; it bleached to *B*. The difference in absorption spectrum before and after bleaching (*A* − *B*) is shown at the right. It has the maximum at about 498 mμ characteristic of regenerated rhodopsin.

Such an experiment is shown in figure 2. Fifteen frog retinas (*Rana pipiens*) were isolated in dim red light, then left in white light for 1½ hours. At the end of this time they were colorless, the retinene$_1$ first formed in bleaching having been reduced to vitamin A$_1$. The retinas were homogenized and preextracted thoroughly with neutral phosphate buffer; this ex-

tract, flooded with retinene$_1$, yields no rhodopsin. Subsequent treatment with 2 per cent digitonin dissolves rhodopsin-protein, which even after losing its carotenoid prosthetic group requires a detergent for its extraction. The rhodopsin-protein was mixed with a digitonin solution of synthetic retinene$_1$, with the result shown at the left of the figure. The extinction at 500 mμ rose rapidly to a maximum value, attained in about 40 minutes at 25°C. On exposure to light this product bleached from *A* to *B*. The difference in absorption spectrum before and after bleaching ("difference spectrum") is shown at the right of the figure. It has the maximum at about 498 mμ characteristic of regenerated rhodopsin. The short-wave-length limb of this difference spectrum is abnormally high, due to some photodecomposition of retinene$_1$ in addition to the bleaching of rhodopsin. A more perfect difference spectrum obtained in a similar experiment is shown in figure 3 (curve 1).[10]

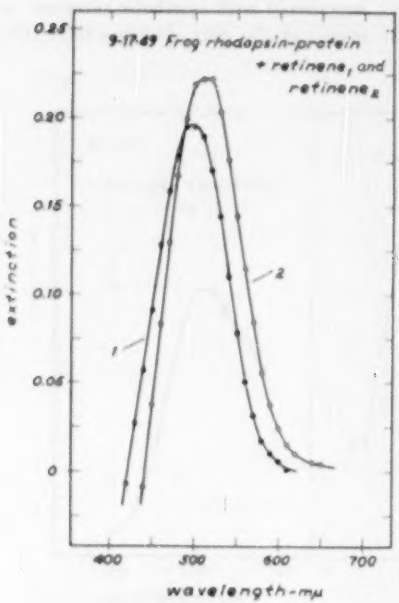


FIGURE 3

Synthesis of light-sensitive pigments from rhodopsin-protein mixed with retinene$_1$ and retinene$_2$. Rhodopsin-protein was extracted with digitonin from frog retinas which had been bleached to colorlessness in daylight. Portions of this were mixed with synthetic retinene$_1$ and retinene$_2$ in digitonin, and let stand in the dark at about 23°C. for 18 hours. The absorption spectra were measured in the dark, then again after bleaching in white light. The differences in the absorption spectra before and after bleaching are shown. That from rhodopsin-protein mixed with retinene$_1$ (1) shows the maximum at about 497 mμ characteristic of synthetic rhodopsin; that from the protein mixed with retinene$_2$ shows a maximum displaced to about 512 mμ.

Do other molecules than rhodopsin-protein and retinene$_1$ take part in this reaction? This question cannot yet be answered with certainty, yet all available evidence suggests that nothing more is needed. As one indication of this, the more highly purified our rhodopsin preparations, the better they regenerate. That shown in figure 1, for example, was made as follows. Cattle retinas were crushed, and the outer segments of the receptor cells

separated from the remaining retinal tissue by differential centrifugation. The outer limbs were tanned in 4 per cent alum, leached with water and neutral phosphate buffer, lyophilized and extracted twice in the dry state with petroleum ether. Finally rhodopsin was extracted with digitonin solution. Few other substances would have emerged from a procedure so narrowly adjusted to the peculiar properties of rhodopsin; and judging from its spectrum this was one of the purest solutions of this pigment yet prepared. Yet when mixed with highly purified retinene$_1$ it yielded an exemplary regeneration.

We have found that rhodopsin can be kept for many days, or dialyzed exhaustively against neutral phosphate buffer, without impairing its capacity to regenerate after bleaching. On the other hand we have not succeeded in increasing regeneration by adding various substances, among them adenosine triphosphate.

We conclude that in all probability this path for the synthesis of rhodopsin involves only the products of its bleaching. Rhodopsin-protein couples with retinene$_1$, and from it synthesizes its prosthetic group. This process probably involves the condensation of two molecules of retinene$_1$.[11] If one wishes to invoke an enzyme in this reaction, it should probably be rhodopsin-protein itself. Yet this view would violate a basic tenet of the enzyme concept, since rhodopsin-protein is removed stoichiometrically by the reaction.

Though these experiments do no more than make very probable that the synthesis of rhodopsin requires only retinene$_1$ and rhodopsin-protein, they make virtually certain that this is a spontaneous—i.e., an exergonic—reaction. It does not have this appearance, since it is a rather complex synthesis. But in fact, it is the bleaching of rhodopsin to retinene$_1$ and protein that requires energy, usually furnished by light. The reversion of retinene$_1$ and protein to rhodopsin is an energy-yielding reaction. This fact is of fundamental importance to an understanding of the visual cycle.

The regeneration of rhodopsin from rhodopsin-protein and retinene$_1$ is inhibited competitively by formaldehyde. One per cent formaldehyde causes an appreciable inhibition, 2–4 per cent almost complete blocking. We interpret this effect as a competition between formaldehyde and retinene$_1$ (vitamin A$_1$ aldehyde) for the amino groups on rhodopsin-protein with which both aldehydes readily couple. The regeneration of rhodopsin is also blocked completely by hydroxylamine (0.15 *M*), which "traps" retinene$_1$, presumably in the form of its oxime.

In every detail so far examined, the rhodopsin system has been closely mimicked by the porphyropsin cycle found in the rods of freshwater vertebrates. It has lately been shown, for example, that retinene reductase taken from either system reduces either retinene$_1$ or retinene$_2$ equally well to the corresponding vitamin A.[6]

Recently several attempts have been made to force the formation of porphyropsin in the eyes of rats[12] and persons[13] by feeding large quantities of vitamin A_2. These seem to have yielded positive effects, though small and still equivocal in meaning.

We have examined the effect of mixing frog rhodopsin-protein with a high concentration of synthetic retinene$_2$. This was made from a highly purified preparation of vitamin A_2[14] by chromatographic oxidation on manganese dioxide.[15] A synthesis of light-sensitive pigment occurs of about the same extent as with retinene$_1$. The difference in absorption spectrum before and after bleaching this product, however, is not characteristic of rhodopsin, but is displaced about 15 mμ toward the red, its maximum lying at about 512 mμ (Fig. 3). This brings it about half-way toward the position characteristic of porphyropsin. Whether we have to deal here with a mixture of photosensitive pigments, or with a modification of spectrum caused by the attachment of the porphyropsin prosthetic group to an abnormal protein, is still to be determined.

We have stressed in this paper the synthesis of photosensitive pigments from the retinenes, since this was accomplished by means employed in the retina. It should be noted, however, that we have in fact carried through for the first time the synthesis of rhodopsin from crystalline vitamin A_1, and the synthesis of a photopigment intermediate between rhodopsin and porphyropsin from pure vitamin A_2. The first step in each of these syntheses—the oxidation of the vitamins A to the corresponding retinenes—was accomplished, not as in the retina, but on manganese dioxide.

Yet we have already described an enzyme in the retina which could perform this oxidation also—retinene reductase, with cozymase (Coenzyme I, DPN) as its coenzyme. A number of comparable enzyme systems are known which reduce aldehydes to alcohols. All of them are potentially reversible, though always, as in this case also, their equilibria lie far over toward the side of reduction, the alcohol. To expose their reversibility it is common practice to use an aldehyde-fixing reagent—dimedone, semicarbazide, hydroxylamine, bisulfite—which by removing the aldehyde promotes a continuous displacement of equilibrium in the oxidative direction.[16]

It is obvious that the reaction described in the present paper, the spontaneous synthesis of rhodopsin, should constitute an efficient process for thus "fixing" retinene$_1$. Coupled with the retinene reductase system, this could drive a continuous though slow oxidation of vitamin A_1 to retinene$_1$, by continuously removing retinene$_1$ to form rhodopsin.

The difficulty with this system as it stands is that the *isolated* retina, though it contains retinene reductase,[8] cozymase and rhodopsin-protein, forms only minimal amounts of rhodopsin from vitamin A_1. Either some additional factor makes this process operate more efficiently in the intact eye; or an alternative path exists for oxidizing vitamin A_1 to retinene$_1$ or to

rhodopsin. We have recently examined both the oxidation of vitamin A_1 to retinene$_1$, and the synthesis of rhodopsin from vitamin A_1, in retinal homogenates and extracts.[17] These experiments will be described in a subsequent paper.

Summary.—Cattle or frog rhodopsin, bleached in solution in the presence of high concentrations of synthetic retinene$_1$, regenerate about 70 per cent of their original content of rhodopsin when replaced in the dark. Rhodopsin is also synthesized *de novo* when its colorless protein moiety is mixed in solution with synthetic retinene$_1$. There is no evidence that other molecules participate in this reaction. It is clearly a spontaneous—i.e., an energy-yielding—process. It is inhibited competitively by formaldehyde, and also with such retinene-trapping reagents as hydroxylamine. When rhodopsin-protein is mixed in solution with synthetic retinene$_2$, a light-sensitive pigment is formed with its spectrum displaced in the direction of porphyropsin.

These experiments in fact accomplish the synthesis of rhodopsin from crystalline vitamin A_1, and of a comparable light-sensitive pigment from highly purified vitamin A_2. The vitamins A were oxidized to the corresponding retinenes, not by a retinal process, but by chromatographic oxidation on manganese dioxide. Nevertheless the pathways of these syntheses may bear a close relation to those which occur in the retina.

* This investigation has been supported in part by a grant from the Medical Sciences Division of the Office of Naval Research.

[1] Kühne, W., "Chemische Vorgänge in der Netzhaut," in L. Hermann, *Handbuch der Physiologie*, F. C. W. Vogel, Leipzig, 1879, vol. 3, part 1, p. 321.

[2] Wald, G., *J. Gen. Physiol.*, **19**, 351 (1935–1936).

[3] Hecht, S., Chase, A. M., Shlaer, S., and Haig, C., *Science*, **84**, 331 (1936).

[4] Chase, A. M., and Smith, E. L., *J. Gen. Physiol.*, **23**, 21 (1939–1940).

[5] Wald, G., Durell, J., and St. George, R. C. C., *Science* (in press) (1950).

[6] Wald, G., *ibid.*, **109**, 482 (1949).

[7] Ball, S., Collins, F. D., Morton, R. A., and Stubbs, A. L., *Nature*, **161**, 424 (1948).

[8] Wald, G., *J. Gen. Physiol.*, **31**, 489 (1947–1948). Based on a reaction discovered by Ball, S., Goodwin, T. W., and Morton, R. A., *Biochem. J.*, **42**, 516 (1948).

[9] The fall in extinction at 500 mμ on adding hydroxylamine to these solutions is caused by the fact that both the native and the added retinene$_1$ they contain is coupled with other molecules, in complexes which still display considerable absorption at 500 mμ. Hydroxylamine destroys these complexes by itself appropriating retinene$_1$, yielding presumably its oxime, with an absorption maximum at about 363 mμ, and almost no absorption at 500 mμ. Hydroxylamine was first used in this laboratory to block the regeneration of rhodopsin, in an experiment by Mr. Jack Durell.

[10] Rhodopsin that has been regenerated or synthesized in solution has its spectrum at slightly lower wavelengths than rhodopsin extracted from the rods. This difference is apparent in the measurements of Chase and Smith,[4] in figure 1 of the present paper, and in recent studies of the regeneration of rhodopsin from early stages of bleaching (Wald et al.[5]; Morton and Collins, Proc. 1st Internat. Congr. Biochemistry, Cambridge, 1949, p. 69). For this reason Morton and Collins have suggested that the rhodopsin regenerated in solution be called *iso*-rhodopsin. No other difference between it and retinal rhodopsin has yet appeared; nor is it yet clear that the regenerated

material is a single molecular species, rather than a mixture of the native and slightly modified pigments.

Note added in proof: Since this account was written we have observed regenerations of 85 per cent in purified cattle rhodopsin, flooded with retinene$_1$, and bleached for 1 minute in intense light.

[11] Wald, G., *Documenta Ophthal.*, **3**, 94 (1949).

[12] Shantz, E. M., Embree, N. D., Hodge, H. C., and Wills, J. H., *J. Biol. Chem.*, **163**, 455 (1946).

[13] Millard, E. B., and McCann, W. S., *J. Applied Physiol.*, **1**, 807 (1949).

[14] I am greatly indebted for this preparation to Dr. Edgar M. Shantz of the Research Laboratories of Distillation Products, Inc., of Rochester, New York.

[15] Wald, G., Meyerhof Festschrift, *Biochem. et Biophys. Acta*, **4**, 215 (1950). Based on a reaction discovered by Morton, R. A., Salah, M. K., and Stubbs, A. L., *Biochem. J.*, **40**, Proc. lix (1946).

[16] Negelein, E., and Wulff, H.-J., *Biochem. Z.*, **293**, 351 (1937). Lutwak-Mann, C., *Biochem. J.*, **32**, 1364 (1938). We owe the basic information on aldehyde fixation to the studies of Carl Neuberg and his coworkers [cf. Neuberg, C., and Kobel, M., *Biochem. Z.*, **188**, 211 (1927)].

[17] Wald, G., and Hubbard, R., PROC. NATL. ACAD. SCI., **36**, 92–102 (1950).

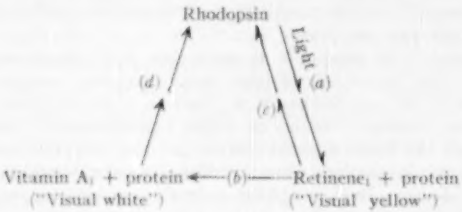
THE SYNTHESIS OF RHODOPSIN FROM VITAMIN A$_1$

BY GEORGE WALD AND RUTH HUBBARD*

BIOLOGICAL LABORATORIES OF HARVARD UNIVERSITY, CAMBRIDGE

Communicated by A. S. Romer, December 30, 1949

A number of years ago it was shown that in the rods of the retina, the light-sensitive pigment rhodopsin takes part in a cycle of the form:[1]



Some progress has been made recently with the chemistry of all these reactions. (*a*) In the bleaching of rhodopsin to retinene$_1$, the light reaction has been isolated and the orange intermediates, lumi- and meta-rhodopsin, characterized.[2] (*b*) Dihydrocozymase (DPN-H$_2$) has been identified as the coenzyme for the reduction of retinene$_1$ to vitamin A$_1$ by retinene reductase.[3] (*c*) Rhodopsin has been synthesized from rhodopsin-protein and retinene$_1$, the latter prepared by oxidizing crystalline vitamin A$_1$ with manganese dioxide.[4]

There remains reaction (*d*), the retinal synthesis of rhodopsin from vitamin A_1. We can identify this with the process which Kühne long ago called "neogenesis," describing it as the formation of rhodopsin from colorless precursors. It occurs appreciably only in the intact eye, and Kühne believed that it requires the cooperation of the pigment epithelium. Since he failed to see any evidence of orange or yellow intermediates in the course of this process, Kühne concluded that it does not retrace the path by which rhodopsin bleaches.[5] Indeed no direct evidence of a reversion of vitamin A_1 to retinene$_1$ can be found ordinarily in the isolated retina.

In the present paper, however, it is shown that isolated retinas, retinal homogenates and aqueous extracts of the retina can all form a little rhodopsin from vitamin A_1. When retinal homogenates are suitably supplemented, they synthesize almost half as much rhodopsin as is regenerated in the living eye during dark adaptation. It is shown also that one mechanism by which the retina can perform this synthesis—though perhaps not the only one—goes over retinene$_1$, and so retraces at least in part the path by which rhodopsin is bleached.

In the recent synthesis of rhodopsin from crystalline vitamin A_1 mentioned above, the first step—the oxidation of vitamin A_1 to retinene$_1$ (vitamin A_1 aldehyde)—was carried out, not by a retinal reaction, but by chromatographic oxidation on manganese dioxide. Retinene$_1$ prepared in this way condenses spontaneously with rhodopsin-protein to form rhodopsin.[4] The success of this procedure suggested that this might be a route for the synthesis of rhodopsin from vitamin A_1 in the retina.

The retina contains a potential mechanism for oxidizing vitamin A_1 to retinene$_1$ in the retinene reductase system. Under equilibrium conditions, this system primarily *reduces* retinene$_1$, transferring two hydrogen atoms to it from dihydro-cozymase:

$$\underset{\text{Retinene}_1}{C_{19}H_{27}CHO} + \text{DPN-H}_2 \quad \underset{\longleftarrow}{\overset{\text{(Retinene reductase)}}{\longrightarrow}} \quad \underset{\text{Vitamin A}_1}{C_{19}H_{27}CH_2OH} + \text{DPN}$$

Similar DPN-enzyme systems in yeast and animal tissues reduce a variety of aldehydes to the corresponding alcohols.[6] All of them are potentially reversible, but always their equilibria lie far over toward the side of reduction. To demonstrate their reversibility it is common practice to "trap" the aldehyde as fast as it is formed, with one of a number of aldehyde-binding reagents. In this way the system is kept from coming to equilibrium, and is driven continuously in the oxidative direction.

By this means one can also oxidize vitamin A_1 to retinene$_1$ in a retinal preparation. To fix the aldehyde, we have used 0.1 *M* hydroxylamine, NH_2OH. This couples spontaneously with retinene$_1$ to form a pale yellow product—presumably retinene$_1$ oxime—with an absorption maximum at

about 355 mμ in hexane and 363 mμ in aqueous solution. The extinction of retinene$_1$ oxime at its maximum is about 1.4 times that of the free aldehyde.

Dark adapted frog retinas (*Rana pipiens*) were isolated, and were bleached in bright light to colorlessness. In this state all their rhodopsin had been converted to vitamin A_1 and protein. They were then homogenized, suspended in neutral phosphate buffer and incubated at room temperature. The untreated homogenate, like the intact retina, does not form observable amounts of retinene$_1$, nor does it do so on adding DPN. In the presence of hydroxylamine, however, the untreated homogenate oxidizes an appreciable fraction of its vitamin A_1 to retinene$_1$; and the yield is greatly increased on adding cozymase.

These observations are illustrated in table 1. A homogenate of 16 frog retinas was divided into 3 equal portions. One was extracted at once with hexane (*a*). The other portions were incubated for 2 hours, one with hydroxylamine (*b*), the other with both hydroxylamine and DPN (*c*), and

TABLE 1

OXIDATION OF VITAMIN A_1 TO RETINENE$_1$ BY A FROG RETINAL HOMOGENATE, INCUBATED FOR 2 HOURS AT 23°C., WITH AND WITHOUT HYDROXYLAMINE AND DPN

The homogenate initially contained vitamin A_1 alone. The table shows the absorption maxima (λ_{max}) of hexane extracts of the homogenate, and the proportions of vitamin A_1 and retinene$_1$ which they contained. The homogenate was suspended in $M/15$ phosphate buffer, pH 6.84, and this buffer was also added in the amounts shown.

SAMPLE	RETINAL HOMOGENATE, ML.	BUFFER, ML.	HYDROXYLAMINE (0.2 M), ML.	DPN, MG.	λ_{max}, mμ	VITAMIN A_1 TRANSFORMED TO RETINENE$_1$, %
a	0.4	1.6	...	...	322	0
b	0.4	0.6	1.0	...	327	28
c	0.4	0.6	1.0	2.0	335	49

were then extracted with hexane. The absorption spectra of these extracts were measured. As shown in table 1, their absorption maxima (λ_{max}) shifted from the initial position characteristic of vitamin A_1 (*a*), about 5 mμ toward the red on incubation with hydroxylamine (*b*), and about 13 mμ toward the red on incubation with both this reagent and DPN (*c*). From this shift of spectrum it could be computed that in (*b*) the final mixture of retinene$_1$ and vitamin A_1 contained about 28 per cent retinene$_1$, in (*c*) about 49 per cent retinene$_1$.

Another experiment is shown in figure 1. Dark adapted retinas were bleached to colorlessness, homogenized with neutral phosphate buffer and 3 equal portions of homogenate were incubated at room temperature with hydroxylamine and DPN. In (*a*) the reaction was stopped at once by adding methyl alcohol to a concentration of 60 per cent; in (*b*) it was stopped after 1 hour, in (*c*) after 2 hours. All three preparations were ex-

tracted with hexane. The absorption spectra of these extracts are shown in the figure. Vitamin A_1 alone was present initially (*a*); about 38 per cent had been converted to retinene$_1$ in one hour (*b*); and about 51 per cent in two hours (*c*). The differences between the initial and later spectra ($b - a$, $c - a$) show that the absorption had risen in the region of the retinene$_1$ maximum (positive differences) and fallen in the region of the vitamin A_1 maximum (negative differences).

It is evident from these data that in the presence of an aldehyde-fixative, retinal preparations can oxidize vitamin A_1 to retinene$_1$. This process is aided by supplementation with DPN, the coenzyme of the retinene reductase system. The added DPN apparently replenishes the supply of coenzyme present in the retina initially, but rapidly destroyed by a nucleotidase after homogenization.[3]

The retina normally possesses a specific process which binds retinene$_1$, its condensation with rhodopsin-protein to form rhodopsin. Like all proper trapping reactions, there is an energy-yielding process. It is not nearly so efficient as the binding of retinene$_1$ by hydroxylamine, and is therefore nearly completely blocked in the presence of this reagent.[4] In the retina, however, rhodopsin-protein may substitute physiologically for hydroxylamine, and may drive a

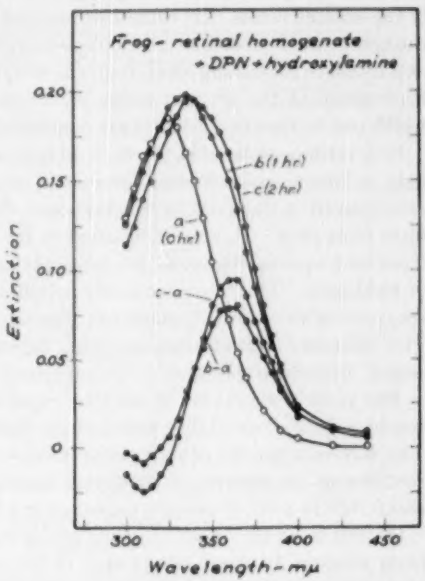


FIGURE 1

Oxidation of vitamin A_1 to retinene$_1$ in a homogenate of frog retinas containing 0.1 *M* hydroxylamine and 2 mg. DPN per ml. In (*a*) the reaction was stopped at once by adding methanol, in (*b*) after 1 hour of incubation at 23°C., in (*c*) after 2 hours incubation. pH 6.8. Absorption spectra of hexane extracts of these preparations are shown (*a*, *b*, *c*). The initial spectrum (*a*) is that of vitamin A_1; but this is displaced toward the red as the reaction proceeds. The changes in absorption ($b - a$, $c - a$) show a rise in the region of the retinene$_1$ maximum (positive differences), a fall in the region of the vitamin A_1 maximum (negative differences). From these data it is computed that 38 per cent of the vitamin A_1 initially present had been oxidized to retinene$_1$ in 1 hour, 51 per cent in 2 hours.

continuous oxidation of vitamin A_1 to retinene$_1$, by removing retinene$_1$ to form rhodopsin.

The difficulty with this idea is that the isolated retina, bleached to colorlessness, does not regenerate rhodopsin efficiently from vitamin A_1, though it contains all the components of the system described above—retinene reductase, cozymase and rhodopsin-protein. It is primarily this behavior of the isolated retina that caused Kühne and later workers to conclude that the synthesis of rhodopsin from colorless precursors requires new materials, and involves something more than the simple reversal of bleaching. This inadequacy of the isolated retina constitutes the primary problem with which our further experiments are concerned.

Frog retinas, isolated in the dark adapted condition, and bleached in the light to colorlessness, contain vitamin A_1 as their only carotenoid. If they are replaced in the dark for several hours, they still appear colorless, or at most faint pink. If, after incubation in the dark, they are extracted with 2 per cent aqueous digitonin, the extract is found to contain a small amount of rhodopsin. This is conveniently revealed by measuring the absorption spectrum of the extract in darkness, then again after bleaching in the light. The difference in absorption spectrum before and after bleaching—the so-called "difference spectrum"—is characteristic of rhodopsin.

The yield of rhodopsin in such an experiment can be estimated by extracting dark adapted frog retinas with digitonin by the same procedure. The difference spectra of such extracts show that on the average each retina contributes an amount of rhodopsin having an extinction at 500 mμ of about 0.08 in 1 ml. of extract, measured in a layer 1 cm. in depth.

On this basis the intact bleached retina replaced in the dark regenerates from vitamin A_1 about 10 per cent of its potential content of rhodopsin. The yield is not increased by adding DPN. This in itself is not significant, since DPN probably does not penetrate the intact tissue.

A homogenate of bleached retinas regenerates about as much rhodopsin in the dark as do intact retinas. When DPN is added to such a homogenate, however, the yield of rhodopsin is approximately doubled.

Such an experiment is shown in figure 2. Sixteen retinas had been bleached in bright light for 45 minutes, and were wholly colorless. They were homogenized with neutral phosphate buffer, and divided into three equal portions. The first of these, extracted at once with hexane, showed the presence of vitamin A_1 alone (A). The other two portions were incubated in the dark at room temperature for 9 hours, one untreated (B), the other with DPN added (C). They were then extracted with digitonin, and the spectra of the extracts measured before and after bleaching. These difference spectra are shown at the left in figure 2. The solid residues from these extractions were reextracted with hexane; the spectra of these extracts are shown at the right.

The hexane extracts show that vitamin A_1 alone is present, both in the original homogenate and in the portions incubated in darkness. The difference spectra at the left are characteristic of rhodopsin. They show that the untreated homogenate regenerated about 10 per cent of rhodopsin, while the homogenate incubated with DPN yielded almost twice this amount, about 18 per cent (cf. table 2).

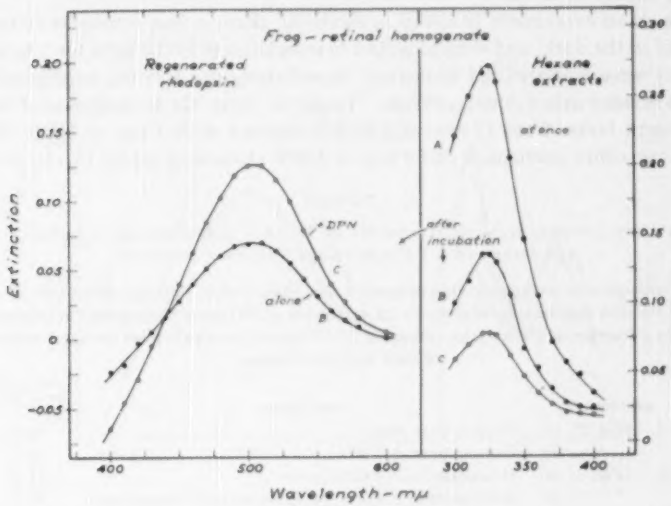


FIGURE 2

Synthesis of rhodopsin from vitamin A_1 in a retinal homogenate, incubated in the dark 9 hours at 23°C., with and without added cozymase (DPN, 4 mg. per ml.). The data at the left show difference spectra of digitonin extracts of these preparations—their differences in absorption before and after exposure to light. They are characteristic of rhodopsin, and show that supplementation with DPN approximately doubles the yield of this pigment. The absorption spectra at the right are of hexane extracts (*A*) of the original homogenate, and (*B*) and (*C*) of the residues of the incubated homogenate after extraction with digitonin. They show that vitamin A_1 is the only carotenoid present initially and at the end of the reaction.

The effectiveness of adding DPN in this synthesis suggests that it was accomplished through the oxidation of vitamin A_1 to retinene$_1$ by the retinene reductase system, retinene$_1$ being trapped by rhodopsin-protein as rapidly as formed. The data show that in the course of this process no free retinene$_1$ accumulates. Kühne's failure to see yellow intermediates during the "neogenesis" of rhodopsin is thus confirmed. Yet this is no assurance that retinene$_1$ is not an intermediate in the reaction.

We have already noted Kühne's insistence upon the co-operation of the pigment epithelium in "neogenesis." We find that when a homogenate of the pigment layers of the eye—pigment epithelium and choroid—is added to a retinal homogenate, the synthesis of rhodopsin from vitamin A_1 is approximately doubled. When both DPN and pigment layers are added to the retinal homogenate, the yield of rhodopsin is again doubled; it now approaches 40 per cent.

Such an experiment is shown in figure 3. Sixteen frog retinas were isolated in the dark, and were bleached to colorlessness in the light for 1 hour. They were homogenized in neutral phosphate buffer and the homogenate was divided into 4 equal portions. To one of these, the homogenate of the pigment layers from 12 eyes was added together with 4 mg. of DPN (*d*). To two other portions, 2 and 4 mg. of DPN alone were added (*b*, *c*). The

TABLE 2

PER CENT REGENERATION OF RHODOPSIN IN RETINAL HOMOGENATES, INCUBATED IN THE DARK FOR 8–9 HOURS UNDER VARIOUS CONDITIONS.

All homogenates were suspended in neutral phosphate buffer, final pH about 6.8. 23°C. The per cent regeneration represents the extinction at 500 mμ of the regenerated rhodopsin as a fraction of the average extinction at 500 mμ of rhodopsin from an equal number of dark adapted retinas.

EXPERIMENT		DESCRIPTION	PER CENT REGENERATION
I.	(Fig. 2) (*a*)	Homogenate alone	10.0
	(*b*)	Homogenate + DPN	18.3
II.	(Fig. 3) (*a*)	Homogenate + DPN	17.9
	(*b*)	Homogenate + DPN + pigment layer homogenate	30.3
III.		Homogenate + DPN	25.4
IV.	(*a*)	Homogenate + pigment layer homogenate	20.3
	(*b*)	Homogenate + pigment layer homogenate + DPN	42.0

fourth portion (*a*) was extracted immediately with hexane; so also was a control portion of pigment layers (*e*). Homogenates *b*, *c* and *d* were incubated for 8 hours in the dark, and were then extracted with digitonin solution. The difference spectra of these extracts are shown at the left of figure 3. After extraction with digitonin, the solid residues were reextracted with hexane; the absorption spectra of these and the control hexane extracts are shown at the right of the figure.

It is clear from figure 3 that the addition of 2 mg. of DPN to a total of 0.7 ml. of reaction mixture produces a maximal effect on the regeneration of rhodopsin (*b*). Doubling this amount does not improve the yield further (*c*). On the other hand the addition of both DPN and homogenized pigment layers approximately doubles the yield of rhodopsin (*d*) (cf. table 2). The pigment epithelium therefore must add something other than DPN to the reacting system.

The spectra of carotenoid extracts at the right of figure 3 show that vitamin A_1 alone is present in the retinal homogenates, vitamin A_1 and xanthophyll in the homogenates containing pigment layers. In no case is any retinene$_1$ apparent, either in the tissues extracted at once or in those incubated in darkness.

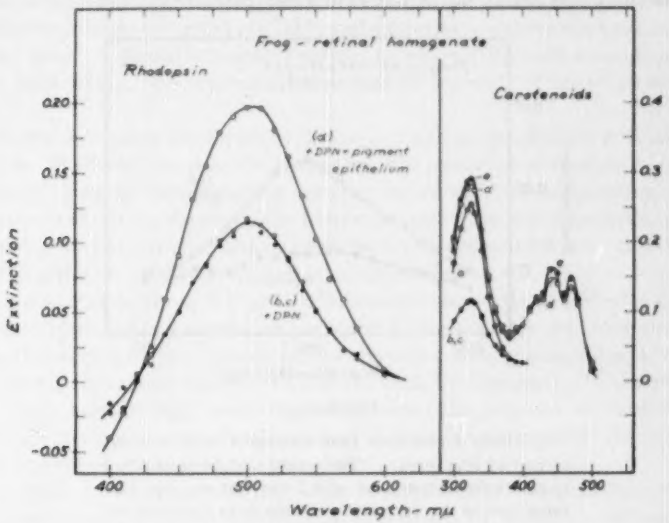


FIGURE 3

Synthesis of rhodopsin from vitamin A_1 in a retinal homogenate, incubated in the dark 8 hours at 23°C., with (*b*) 2 mg. and (*c*) 4 mg. DPN added; and with (*d*) 4 mg. DPN and pigment layer homogenate added. The data at the left show the difference spectra of digitonin extracts of these preparations. They are characteristic of rhodopsin, and show that supplementation with 2 mg. DPN is as effective as 4 mg. DPN (*b*, *c*); but that the addition of pigment layer homogenate approximately doubles the yield of rhodopsin. The absorption spectra at the right are of hexane extracts of the solid residues of these preparations, and of control portions of the original retinal homogenate (*a*) and of the pigment layer homogenate (*e*). They show that the retinal preparations contain vitamin A_1, the pigment layer preparations vitamin A_1 and xanthophyll as their only carotenoids at the beginning and end of the reaction. Vitamin A_1 is responsible for the single absorption band in the ultra-violet at about 325 mμ, xanthophyll for the complex spectrum in the visible region.

The yields of rhodopsin obtained in retinal homogenates by all the types of treatment described above are summarized in table 2. We have, however, also observed the synthesis of rhodopsin from vitamin A_1 in aqueous solution.

An aqueous digitonin extract of bleached frog retinas contains rhodopsin-protein and retinene reductase. If freshly prepared, it also contains cozymase, though this is destroyed within a few hours by the nucleotidase which is present.[3] The addition of a high concentration of vitamin A_1 to such an extract might be expected to force the production of some retinene$_1$, and from this could induce the synthesis of some rhodopsin.

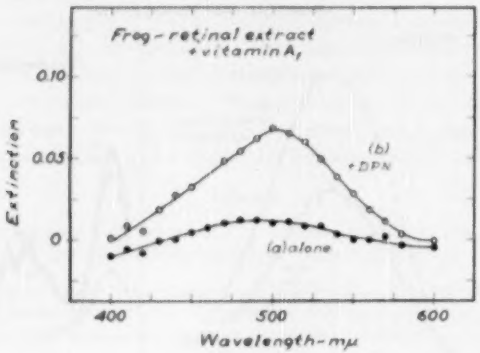


FIGURE 4

Synthesis of rhodopsin from vitamin A_1 in an aqueous extract of frog retinas. The retinas had been bleached to colorlessness, extracted with 2 per cent aqueous digitonin, and to this crystalline vitamin A_1 in digitonin was added. Half of this preparation was incubated in the dark overnight at 23°C. without further treatment (*a*), the other half with 4 mg. DPN added per ml. of extract (*b*). The data show the difference spectra of the final solutions. The solution without added DPN had synthesized a scarcely perceptible amount of rhodopsin; that supplemented with DPN had regenerated about 10 per cent as much rhodopsin as was present originally in the dark adapted retinas from which it was prepared.

Such an experiment is shown in figure 4. Twelve dark adapted frog retinas were isolated, bleached to colorlessness, homogenized and the homogenate extracted for 3 hours with digitonin. The extract was cleared by centrifuging at high speed, and to it a clear solution of crystalline vitamin A_1 in digitonin was added. The mixture was divided into halves, and to one portion 3.5 mg. of DPN was added. Both portions were left in the dark for 12 hours, and then their difference spectra were measured. These are shown in the figure. The solution to which no DPN had been added (*a*) regenerated an almost negligible amount of rhodopsin. That supplemented with DPN (*b*) had regenerated about 10 per cent as much rhodopsin as was

present originally in the dark adapted retinas from which the solution was prepared.

In summary, therefore, we have found that rhodopsin is synthesized from vitamin A_1 in intact retinas and retinal homogenates, with yields of about 10 per cent. In retinal homogenates supplemented with either DPN or pigment layer homogenate, the yields are about 20 per cent; and in homogenates supplemented with both DPN and pigment layers they rise to about 40 per cent. Aqueous extracts of retina, to which DPN and vitamin A_1 have been added, also synthesize rhodopsin with a yield of about 10 per cent.

The mechanism of rhodopsin synthesis in these preparations is still uncertain. We know one possible mechanism, the oxidation of vitamin A_1 to retinene$_1$ by the retinene reductase system, coupled with the condensation of retinene$_1$ with rhodopsin-protein to form rhodopsin. Yet this apparatus, as we find it in the isolated retina, appears unable to account for more than a small fraction of the rhodopsin formed in the intact eye.

It seems reasonably clear that DPN stimulates the synthesis of rhodopsin in our preparations by supplying coenzyme to the retinene reductase system. Pigment layer homogenate may act upon the same system, by supplying respiratory factors which drive it in the oxidative direction. In this regard it is perhaps significant that riboflavine, the common oxidant of DPN-H_2 in cellular respiration, is present in high concentration in the pigment epithelium, though little is found in the retina.[7]

We think it almost certain that some rhodopsin is synthesized through the intermediate oxidation of vitamin A_1 to retinene$_1$ by retinene reductase. It is conceivable that by coupling with respiratory mechanisms, the efficiency of the retinene reductase system is so greatly increased that it can account for the whole synthesis of rhodopsin from vitamin A_1. On the other hand alternative pathways for this process may exist, and may even be of major importance. These problems are being investigated further.

Summary.—Intact frog retinas and retinal homogenates can synthesize rhodopsin from vitamin A_1 in amounts about 10 per cent as great as are formed during dark adaptation *in vivo*. When either cozymase (DPN) or a homogenate of the pigment layers of the eye—pigment epithelium and choroid—are added to a retinal homogenate, the yield of rhodopsin is approximately doubled; and by the addition of both DPN and pigment layer homogenate, the yield is doubled again, bringing it to about 40 per cent. Aqueous extracts of retina, supplemented with DPN and vitamin A_1, also synthesize rhodopsin with a yield of about 10 per cent.

The mechanism of this synthesis is not yet wholly understood. Retinal homogenates slowly oxidize vitamin A_1 to retinene$_1$ (vitamin A_1 aldehyde) in the presence of such an aldehyde-binding reagent as hydroxylamine. This process, like rhodopsin synthesis, is aided by supplementation with

DPN, the coenzyme of the retinene reductase system. Hydroxylamine apparently drives the retinene reductase system in the oxidative direction, by removing retinene$_1$ as fast as it is formed.

In the retina, rhodopsin-protein may act similarly, since it is known to bind retinene$_1$ by condensing with it spontaneously to form rhodopsin. This is almost surely one mechanism for the synthesis of rhodopsin from vitamin A_1. Yet the isolated retina, which contains all the components of this system, regenerates very little rhodopsin from vitamin A_1. Either the retinene reductase system operates much more efficiently in the whole eye than in the isolated retina, through the action of such auxiliary factors as may be added by the pigment epithelium; or some rhodopsin is synthesized from vitamin A_1 by alternative mechanisms still to be explored.

* This investigation was supported in part by a grant from the Medical Sciences Division of the Office of Naval Research.

[1] Wald, G., *J. Gen. Physiol.*, **19**, 351, 781 (1935–1936).

[2] Wald, G., Durell, J., and St. George, R. C. C., *Science*, in press (1950).

[3] Wald, G., and Hubbard, R., *J. Gen. Physiol.*, **32**, 367 (1948–1949). Wald, G., *Science*, **109**, 482 (1949).

[4] Wald, G., and Brown, P. K., PROC. NATL. ACAD. SCI., **36**, 84–92 (1950).

[5] Kühne, W., "Chemische Vorgänge in der Netzhaut," in L. Hermann, *Handbuch der Physiologie*, F. C. W. Vogel, Leipzig, 1879, vol. 3, part 1, p. 317.

[6] Negelein, E., and Wulff, H.-J., *Biochem. Z.*, **293**, 351 (1937). Lutwak-Mann, C., *Biochem. J.*, **32**, 1364 (1938).

[7] von Euler, H., and Adler, E., *Ark. Kemi, Mineral., Geol.*, **11B**, No. 21 (1933); *Z. physiol. Chem.*, **228**, 1 (1934). Wald, G., *J. Gen. Physiol.*, **19**, 781 (1935–1936).

---

*ANTIBIOTIC SUBSTANCES FROM BASIDIOMYCETES. VI. AGROCYBE DURA**

BY FREDERICK KAVANAGH, ANNETTE HERVEY AND WILLIAM J. ROBBINS

DEPARTMENT OF BOTANY, COLUMBIA UNIVERSITY, AND THE NEW YORK BOTANICAL GARDEN

Communicated December 19, 1949

In a previous report[1] from this laboratory, *Agrocybe dura* (L386.10)[2] was found to evidence considerable activity against *Staphylococcus aureus* (H) and *Escherichia coli* when tested by the streak of disk methods. Culture liquids of this fungus were found to have some activity. Further investigation has resulted in additional information on the antibiotic properties of *Agrocybe dura*.

Antibiotic Material in Liquid Culture.—Culture liquids with antibacterial activity were produced by growing the fungus at 25°C. in 2800 ml. Fernbach flasks containing a corn steep medium, as previously described.[3]

The fungus grew with medium rapidity; it required from 6 to 8 weeks to produce culture liquids of from 256 to 512 dilution units per ml. against *Staph. aureus*. When the active culture liquid was decanted and replaced with a liter of fresh corn steep medium, an activity of 256 dilution units per ml. against *Staph. aureus* developed in about two weeks.

A sample of culture liquid was Seitz filtered at its natural pH of 4.6, at pH 3.0, and at pH 7.0. More than three-fourths of the activity was removed by the filtration at pH 3.0; about one-half at pH 4.6; and none at pH 7.0. Adjusting to pH 3.0 and then neutralizing did not destroy the active substance. The filtered solutions of pH 4.6 and pH 7.0 were placed at 11°C. for 26 days with assay at frequent intervals. Neither solution lost activity.

Concentration of Active Material.—Preliminary experiments showed that culture liquid dried in the air lost about one-half of its activity. Because of this observation, which was confirmed by later experiments, care was used to avoid drying at any stage in the concentration of the active material.

Aqueous concentrates of antibacterial material were prepared by extracting the culture liquid with methyl-iso-butyl ketone. The solvent was evaporated to small volume *in vacuo*. The residue in the still was dissolved in alcohol which was diluted with water or the residue was dissolved in ether which was then washed with bicarbonate solution as described later and evaporated over a layer of water. Such concentrates had an activity of 1,000,000 or more dilution units per mg. against *Staph. aureus*.

Isolation of Crystalline Material.—A crystalline antibiotic substance was isolated from the culture liquid. The procedure for one lot was as follows: A volume of 10.5 liters of culture liquid with an activity of 128 dilution units per ml. against *Staph. aureus* was extracted twice with 750 ml. of methyl-iso-butyl-ketone; the ketone was evaporated to very small volume *in vacuo;* and ether added. The ether solution was extracted with 2% sodium bicarbonate solution to remove acids and part of the colored substances. Brown colored crystals were obtained upon evaporation of the ether at room temperature. These crystals were purified by recrystallization from 20% alcohol or from ether.

Characteristics of Crystalline Material.—The white crystals obtained were stable in air for about one day after which they turned black and became insoluble in ether. The crystals did not change in color when kept at 4°C. *in vacuo*. On attempting a melting point in air, the crystals turned black and then exploded at a temperature of about 145°. In a capillary *in vacuo*, the crystals decomposed at about 90° and evolved gas at 140°C. An analysis[4] indicated 65.58% C and 4.31% H. We propose to name this substance agrocybin.

Agrocybin was soluble in alcohol, acetone, ether, chloroform, methyl-

iso-butyl-ketone, slightly soluble in water and insoluble in hexane. The optical rotation of agrocybin in alcohol was zero for the sodium D line. The ultra-violet absorption spectrum of agrocybin in alcohol had peaks at 216, 224, 269, 286, 304 and 325 mμ. The curve obtained in pH 6.0 phosphate buffer was nearly the same as that obtained in alcoholic solution. The curve, figure 1, resembled greatly that of nemotin A.[5] There were sufficient differences, however, to indicate that the two substances were not the same.

Agrocybin was inactivated by heating to 100° at pH 8.5 and higher but not at pH 6.5. It was not inactivated in one hour at 25° and pH between 3.0 and 11.0. It did not distill with steam. Agrocybin was not extracted from ether by aqueous solution of pH 8.5 and probably is a neutral or a very weakly acidic substance.

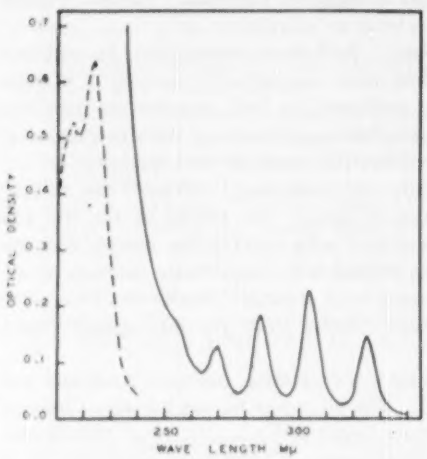


FIGURE 1

Absorption spectrum of crystalline agrocybin in alcohol. The optical density is given for solutions measured in a 1 cm. cuvette. The dashed line refers to a concentration of 1 microgram per milliliter, and the continuous line to a concentration of 10 micrograms per milliliter.

Evidence for Existence of Agrocybin in an Inactive Form.—Part of the activity of a culture liquid was in an inactive ketone-insoluble form which was made active and ketone-soluble by boiling the solution.[6] In this respect the behavior of the culture liquid of *Agrocybe dura* was similar to that found previously[7] for the culture liquid of *Polyporus biformis*. A concentrate of the active substance free in the culture solution and a concentrate prepared from a culture solution heated after the free substance had been removed had the same antibacterial and antiluminescent activities. It is probable that the antibacterial substance formed by heating the culture liquid was the same as that which occurred free in the liquid. Since the antibacterial titer of a culture liquid was doubled by heating there appeared to be as much of the active substance in the inactive form as there was free.

Antibacterial and Antiluminescent Activity.—The antibacterial and antiluminescent activites of agrocybin were obtained by the methods in use in this laboratory[8] with the results set forth below. The activity is expressed as the minimum inhibitory concentration in micrograms per ml.

BACTERIA	ACTIVITY	BACTERIA	ACTIVITY
Bacillus mycoides	0.5	*Mycobacterium smegma*	1
Bacillus subtilis	0.5	*Photobacterium fischeri*	0.125
Escherichia coli	1	*Pseudomonas aeruginosa*	0.5
Klebsiella pneumoniae	1	*Staphylococcus aureus* (H)	1
Mycobacterium phlei	0.5		

The antiluminescent activity against *Photobacterium fisheri* is given in the following table:

TIME OF INCUBATION	ACTIVITY
10 min.	0.016
1 hrs.	0.008
2 "	0.004
3 "	0.002
24 "	0.064

Agrocybin belongs in the group of substances inhibitory for gram-positive, gram-negative and acid-fast bacteria. The antibacterial activities resemble those of biformin except that agrocybin is much more active than biformin against *B. mycoides* and *P. aeruginosa*. Although there may be a close chemical relation between agrocybin and nemotin A, as indicated by similar ultra-violet spectra, agrocybin is much more active against the gram-negative bacteria than is nemotin A.

The antiluminescent activity of agrocybin is much higher than that of any other substance tested and is about 60 times as great as the best preparation of biformin,[7] the second most active substance investigated.

Antifungal Activity.—The activity of two concentrates prepared from the culture liquid was measured against nine fungi by serial dilution[8] and by an agar disk method. The fungi were: *Aspergillus niger, Chaetomium globosum* (USDA 1042.4), *Gliomastix convoluta* (PQMD4c), *Memnoniella echinata* (PQMD1c), *Myrothecium verrucaria* (USDA 1334.2), *Penicillum notatum* (832), *Phycomyces Blakesleeanus* (+ strain), *Stemphylium consortiale* (PQMD41b) *and Trichophyton mentagrophytes.*

When measured by the serial dilution method, the concentrates were about one-eighth to one-fourth as active against the fungi as against *Staph. aureus* except for *P. Blakesleeanus* and *G. convoluta* which were 32 and 128 times respectively as resistant as *Staph. aureus.*

Activity in the Presence of Blood.—Agrocybin and the antibacterial material in the culture liquid and in concentrates were inactivated to a considerable degree by rabbit and by human blood. The activity of agrocybin against *Staph. aureus* was reduced to $^1/_{32}$ by incubation for 3 hours in a beef extract medium containing 5% whole human blood. A ketone concentrate was not reduced in activity by incubation with 5% human plasma and lost only one-half of its activity when incubated with 5% washed or hemolyzed human red cells.

Animal Toxicity.—A solution of 1 mg. per ml. of agrocybin was prepared in 0.7% sodium chloride solution and quantities ranging from 0.1 ml. to 1 ml. were injected into a tail vein of 16–18 g. Carworth Farms CFI male white mice. The larger the dose, the sooner the mice died. The animals receiving 1 mg. of agrocybin died within 1 hour, those receiving 0.1 mg. died in about 18 hours. The LD_{50} was less than 6 mg. per kg. body weight of white mice. Agrocybin caused a dermatitis in a human subject susceptible to bioformin.

Because of its high toxicity and inactivation by blood, experiments on possible therapeutic action were not attempted.

* This investigation was supported in part by grants from The Commonwealth Fund and The Lillia Babbitt Hyde Foundation.

[1] Hervey, A. H., *Bull. Torrey Bot. Club*, **74**, 476–503 (1947).

[2] We are indebted to Dr. José Emilio Santos Pinto Lopes, Instituto Botanico, Lisbon, Portugal, for the culture of *Agrocybe dura*.

[3] Robbins, W. J., Kavanagh, F., and Hervey, A., these PROCEEDINGS, **33**, 171–176 (1947).

[4] Analysis by Mr. Joseph F. Alicino.

[5] Kavanagh, F., Hervey, A., and Robbins, W. J., these PROCEEDINGS, in press.

[6] Some evidence indicated that boiling not only transformed inactive into active material but destroyed part of the active material. The antibacterial titer of a boiled culture liquid was the net of these two opposite effects.

[7] Robbins, W. J., Kavanagh, F., and Hervey, A., these PROCEEDINGS, **33**, 176–182 (1947).

[8] Kavanagh, F., *Bull. Torrey Bot. Club*, **74**, 303–320, 414–425 (1947).

[9] Anchel, M., Hervey, A., Kavanagh, F., Polatnick, J., and Robbins, W. J., these PROCEEDINGS, **34**, 498–502 (1948).

---

*QUANTITATIVE EFFECTS OF PROTEIN ENRICHMENT OF DIET UPON GROWTH AND EARLY ADULT LIFE**

BY HENRY C. SHERMAN, CONSTANCE S. PEARSON AND MARY E. BAL

DEPARTMENT OF CHEMISTRY, COLUMBIA UNIVERSITY

Communicated December 10, 1949

Previous publications have shown that with experimental animals (rats) eating *ad libitum* of a basal diet, adequate for their normal nutrition but of relatively low calcium content (0.18 to 0.2 per cent of the air-dry food mixture) the weight of the body may grow more rapidly than does the weight of calcium which the body contains. In such cases, notwithstanding continuously positive calcium balances, there may develop a low-calcium condition of body, shown by a dip in the curve which represents the *percentage* of calcium in the growing body.[1]

Moreover, among subjects of the same age and sex, living on the same

TABLE 1

FINAL AVERAGE RECORDS FOR RATS FED A WHEAT-AND-MILK BASAL DIET, ALONE AND WITH ADDED PROTEIN AS POULTRY MEAT

(10 Rats of Each Sex on Each Point, Full Litter-Mate Control)

	ON DIET 16 (14% PROTEIN); BASAL DIET OF WHEAT AND MILK MEAN WITH ITS P.E.	C.V.	ON DIET 16 P 5 (24% PROTEIN); BASAL DIET + MEAT MEAN WITH ITS P.E.	C.V.	DIFFERENCE WITH ITS P.E.	CRITICAL RATIO
Gain in weight, 28–56 days of age						
Females, g.	39.7 ± 1.59	19	61.6 ± 2.05	16	21.9 ± 2.59	8.5
Males, g.	48.5 ± 2.71	26	81.5 ± 2.51	14	33.0 ± 3.69	8.9
Age of females at birth of first young, days	132 ± 6.3	22	98 ± 3.0	14	34 ± 7.0	4.9
Young borne per female	41 ± 3.7	42	44 ± 4.0	43	3 ± 5.4	0.6
Young reared per female	29 ± 2.5	41	35 ± 3.5	47	6 ± 4.3	1.4
Average weight of young at 28 days of age, g.	40.5 ± 0.33	21	45.3 ± 0.32	20	4.8 ± 0.46	10.4
Length of reproductive life of females, days	368 ± 31.7	40	349 ± 20.5	28	19 ± 38	0.5
Reproductive life as per cent of life cycle, %	44 ± 3.9	42	47 ± 2.3	23	3 ± 4.6	0.7
Length of life						
Females, days	848 ± 22.3	12	764 ± 40.5	25	84 ± 46	1.8
Males, days	741 ± 21.0	13	748 ± 23.4	15	7 ± 31	0.2

TABLE 2

ALL AVAILABLE CASES FROM DIET 16 P 5 AND MOST CLOSELY RELATED CONTROLS FROM DIET 16

	DIET 16 MEAN WITH ITS P.E.	C.V.	DIET 16 P 5 MEAN WITH ITS P.E.	C.V.	DIFFERENCE WITH ITS P.E.	CRITICAL RATIO
Females:						
Gain in weight, 28–56 days, g.	(26)* 42 ± 1.2	22	(28) 61 ± 1.3	17	19 ± 1.7	11.2
Age at birth of first young, days	(26) 122 ± 3.33	21	(28) 106 ± 2.85	21	16 ± 4.4	3.6
Weight at 100 days, g.	(12)† 141 ± 2.77	10	(20) 168 ± 2.32	9	27 ± 3.6	7.5
Weight at 210 days, g.	(14) 189 ± 2.02	6	(17) 206 ± 2.63	8	17 ± 3.3	5.2
Weight at 300 days, g.	(20) 205 ± 2.83	9	(17) 225 ± 3.90	11	20 ± 4.8	4.2
Males:						
Gain in weight, 28–56 days, g.	(20) 51 ± 1.9	25	(20) 80 ± 1.6	15	29 ± 2.0	14.5
Weight at 100 days, g.	(20) 186 ± 3.45	12	(20) 218 ± 3.77	11	32 ± 5.1	6.3
Weight at 210 days, g.	(20) 292 ± 3.80	9	(20) 309 ± 5.52	12	17 ± 6.7	2.5
Weight at 300 days, g.	(20) 308 ± 3.63	8	(20) 327 ± 5.28	11	19 ± 6.3	3.0

* Figures in parentheses indicate number of cases.

† Numbers vary due to exclusion of pregnant females, and to deaths.

diet, those showing more rapid growth in weight as an individual characteristic tend to show the temporary low-calcium condition of body in a more pronounced degree.[2]

Still further, it is found that when growth is accentuated by adding protein—alone or as meat—to the already adequate basal diet, this tends to induce a more marked low-calcium condition with symptoms suggestive of those of calcium deficiency.[3] While in the majority of such cases there is spontaneous recovery, some of the females which had been thus stimulated to rapid growth and early puberty then broke down in reproduction.

The experimentation has therefore been continued so that our total findings may be given a more quantitative character, the reporting of which is the purpose of the present paper.

Of 40 young females (rats) on the basal diet (Laboratory Diet No. 16, or Diet A), 3 or 7.5 per cent died between the ages of 1 month ("end of infancy") and 1 year (early adulthood); while of 40 of the same age, sex, heredity and nutritional background, but receiving a supplement of lean meat in addition to the basal diet, 7 or 17.5 per cent died during the same age period. The protein supplement in these cases was 5 (or in one series, 10) grams of lean meat on each of six days each week. This increased the protein of the diet from about 14 per cent to about 24 per cent of the dry matter of the food.

Quantitative Comparison of Biochemical Performance in the Life History.—Findings on several points which have proved noteworthy criteria in the judgment of biochemical well-being and nutritional performance are shown quantitatively in tables 1 and 2.

The data summarized in table 1 show a marked increase of gain in body weight when the basal diet of wheat and milk is supplemented with meat as described above. The difference with females was 8.5, and with males 8.9, times its Probable Error. These Critical Ratios are very clearly significant. This is strongly confirmed by the corresponding data in table 2 which includes the means of all available cases of experimental animals on the protein-supplemented diet (Laboratory Diet 16 P 5) and of all the most closely related—not always litter-mate—controls on Diet 16. Here the increased growth due to the addition of meat to the basal wheat and milk diet (Diet 16) was 11.2 times its Probable Error with females, and 14.5 times with males. Very clearly the findings of the two series of experiments show that the difference in diet resulted in a difference in growth which was undoubtedly statistically significant.

The protein supplement which thus increased the rate of growth appears also to have stimulated an earlier development of puberty as shown by earlier reproduction when the sexes grew up together. The observed difference is 4.9 times its Probable Error in the series of 10 females with 10 litter-mate controls; and 3.6 times in the larger series with 28 experi-

mental females and 26 controls, all closely related though not all littermate controls. Consistently with the usual interpretation in this field of research, we judge these differences to be statistically significant.

The numbers borne, and also the numbers reared, per female, were essentially the same on each of the two diets here compared. The average weight of offspring at the age of 28 days (conventionally considered the end of infancy in the rat) was distinctly higher in the families receiving the supplementary meat ration. The latter, however, seems to have had no clearly significant influence upon the duration of reproductive life as measured in the females or upon the lengths of life in the animals represented in table 1. Some of the additional animals included in the data of table 2 are still living.

Consistently with their greater gains in early growth, the rats fed supplementary protein showed distinctly higher body weights at 100 days of age, but the differences were less pronounced at 210 and at 300 days.

Body weights at higher ages, average lengths of life of larger numbers, and the influence of combined supplementation of the diet with protein and calcium are being studied in experiments still in progress.

* Aided by grants from the John and Mary R. Markle Foundation, and the Williams-Waterman Fund of the Research Corporation, to Columbia University.

[1] Sherman, H. C., Campbell, H. L., and Lanford, C. S., PROC. NATL. ACAD. SCI., **25**, 16–20 (1939).

[2] Lanford, C. S., Campbell, H. L., and Sherman, H. C., *J. Biol. Chem.*, **137**, 627–634 (1941).

[3] Sherman, H. C., and Pearson, C. S., PROC. NATL. ACAD. SCI., **33**, 264–266, 312–314 (1947); **34**, 585–587 (1948). Sherman, H. C., Ragan, M. S., and Bal, M. E., *Ibid.*, **33**, 356–358 (1947).

DISTRIBUTION OF PRESSURE ON CURVED PROFILES IN SUPERSONIC GAS FLOW WITH VARIABLE ENTROPY

BY T. Y. THOMAS

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY

Communicated December 24, 1949

1. Introduction.—A theory of pressure distribution on curved profiles immersed in a gas moving at supersonic speed was presented in a previous note in these PROCEEDINGS and pressures were calculated, to first approximations, for a circular arc profile based on this theory.[1] The formal basis of the theory involves the representation of pressure as a function $p(\omega, S)$, where ω denotes the inclination of the stream lines and S the entropy of the gas, and the power series expansion of $p(\omega, S)$ about a point of inclination ω on the profile. It was pointed out that the second approximation of pres-

sure required the integration of a more or less complicated differential equation of the first order and that this integration could best be performed by use of a high-speed computing machine which was not immediately available to us.

The discussion of the following note[2] is based on the possibility of expansion of the function $p(\omega, S)$ about points of the shock line assumed to originate at the vertex of the profile. Thus we consider the series

$$p(\omega, S) = p(\omega, S_*) + \frac{\partial p(\omega, S_*)}{\partial S}(S - S_*) + \frac{1}{2}\frac{\partial^2 p(\omega, S_*)}{\partial S^2}(S - S_*)^2 + \ldots$$

in which ω is the inclination of the stream line at a point immediately behind the shock line and S_* denotes the entropy at this point. Taking the arbitrary constant in the definition of the entropy function to be such that $S = 0$ for points on the profile, and denoting by $p = \theta(\omega)$ the pressure distribution over the profile, the above equation gives

$$\theta(\omega) = p(\omega, S_*) - \frac{\partial p(\omega, S_*)}{\partial S} S_* + \frac{1}{2}\frac{\partial^2 p(\omega, S_*)}{\partial S^2} S^2 - \ldots.$$

Termination of this series after its first term gives the relation $\theta(\omega) = p(\omega, S_*)$ used for the first approximation of pressure on the profile. Including the second term we have the relation

$$\theta(\omega) = p(\omega, S_*) - \frac{\partial p(\omega, S_*)}{\partial S} S_*. \qquad (1)$$

Under the assumption that the flow is uniform in front of the shock line, the quantity S_* is an explicit function of ω alone. In §3 it will be shown that the derivative $\partial p(\omega, S_*)/\partial S$ is likewise given explicitly as a function of ω in view of the fact that the curvature K of the stream lines, which is naturally introduced by the process of differentiation, cancels from the several terms which make up the expression for this derivative. Hence it is feasible to determine the pressure along the profile from (1) by means of an ordinary computing machine.

2. *Résumé of Formulas for First Approximation of Pressure.*—By the first approximation formula the pressure at any point of the profile, at which the inclination is ω, is given by the function which expresses the pressure in terms of the inclination ω of the stream lines immediately behind the shock line. Denoting by p_1 the constant pressure of the uniform flow in front of the shock line we have

$$\frac{p(\omega, S_*)}{p_1} = \frac{2\gamma}{\gamma + 1}(M^2 \sin^2 \alpha - 1) + 1, \qquad (2)$$

where M is the Mach Number of the uniform flow and $\gamma = 1.405$ for air;

also α is the inclination of the shock line and is related to the inclination ω of the stream lines by the formula

$$\tan \omega = \frac{2[(M^2 - 1)\tan^2 \alpha - 1]}{[(\gamma - 1)M^2 + 2]\tan^2 \alpha + [(\gamma + 1)M^2 + 2]\tan \alpha}. \quad (3)$$

Formulas (2) and (3) permit the determination of the first approximation of pressure. While the quantity S_* does not enter into this approximation, we note that it is given by $S_*(\omega) = Jc_v \log \Delta(\omega)$ where

$$\Delta(\omega) = \frac{(p/p_1)/(\rho/\rho_1)^\gamma}{(p_v/p_1)/(\rho_v/\rho_1)^\gamma}, \quad (4)$$

with

$$\left.\begin{aligned} \frac{p}{p_1} &= \frac{2\gamma}{\gamma + 1}(M^2 \sin^2 \alpha - 1) + 1, \\ \frac{\rho}{\rho_1} &= \frac{2(M^2 \sin^2 \alpha - 1)}{(\gamma - 1)M^2 \sin^2 \alpha + 2} + 1. \end{aligned}\right\} \quad (5)$$

Here J and c_v have the usual meaning, ρ_1 is the density in the uniform flow, while p and ρ are pressure and density at points immediately behind the shock line. Also p_v and ρ_v in (4) stand for pressure and density behind the shock line and at the vertex V of the profile; these quantities are therefore determined by (5) in which α is the inclination of the shock line at V.

3. Derivation of the Formulas for the Derivatives of the Function $p(\omega, S)$ Along the Shock Line.—To determine the derivative $\partial p/\partial S$ in (1) we first consider the relations[3]

$$\frac{\partial p}{\partial \omega} = \frac{dp}{d\sigma}\Big/\frac{d\omega}{d\sigma} = \frac{dp/d\sigma}{K}, \quad (6)$$

$$\frac{dp}{d\sigma} = p_{1\alpha}\frac{dx^\alpha}{d\sigma} = -\rho A_{\alpha\beta}\zeta^1_\alpha \frac{dx^\alpha}{d\sigma}, \quad (7)$$

at points immediately behind the shock line, where σ denotes arc length along the stream lines. But from the definition of the ζ^i_α it readily follows that $\zeta^1_\alpha\, dx^\alpha/d\sigma = 0$ and $\zeta^2_\alpha\, dx^\alpha/d\sigma = 1/v$, where v is the magnitude of the velocity. Hence from (7) we have

$$\frac{dp}{d\sigma} = -\rho A_{\alpha\beta}\zeta^2_\alpha \frac{dx^\alpha}{d\sigma} = -\rho \frac{A_{\alpha\beta}}{v}, \text{ or}$$

$$\frac{dp}{d\sigma} = -\rho \frac{B_{\alpha\beta}\kappa}{v} = \frac{-\rho B_{\alpha\beta}K}{vG_0(M, \alpha)}, \quad (8)$$

where we have put $A_{ij} = B_{ij}\kappa$, $K = G_0\kappa$ with κ the curvature of the shock line. But[3]

$$G_0 = -\frac{B_{12}}{vu_n} + \frac{B_{22}u_t}{v^2u_n}$$

in which $B_{12} = 4u_nu_t/\gamma + 1$; here u_t denotes the tangential component and u_n the normal component of velocity behind the shock line. Hence

$$\frac{B_{22}}{vG_0} = vu_n\left[\frac{v}{u_t} + \frac{4}{(\gamma + 1)G_0}\right]$$

and making this substitution in (8) we have

$$\frac{dp}{d\sigma} = -\rho vu_nK\left[\frac{v}{u_t} + \frac{4}{(\gamma + 1)G_0}\right] \tag{9}$$

along the shock line. Finally from (6) and (9) we find

$$\frac{\partial p}{\partial \omega} = -\rho vu_n\left[\frac{v}{u_t} + \frac{4}{(\gamma + 1)G_0}\right]. \tag{10}$$

It is to be observed from (10) that the derivative $\partial p/\partial\omega$ along the shock line is independent of the curvatures of the stream lines or shock line and in fact depends only on the inclination of the shock line for given uniform incident flow.

The formula for the derivative $\partial p/\partial S$ in (1) is now obtained by combining (10) with the following relation[4]

$$\left[(m^2 - 1)\left(\frac{\partial p}{\partial \omega}\right)^2 - \rho^2v^4\right]\frac{d\omega}{d\sigma} = -\tau A\rho^2v^3\frac{\partial p}{\partial S} \tag{11}$$

in which $\tau = (\gamma - 1)/c_0$, m is the variable Mach Number for flow behind the shock line, and A is a certain quantity which is constant along stream lines. At points immediately behind the shock line A is given by[5]

$$A = \frac{-[\rho]^2u_1\kappa}{\rho_1\rho p} = \frac{-[\rho]^2u_1K}{\rho_1\rho pG_0} \tag{12}$$

in which $[\rho] = \rho - \rho_1$ stands for the difference in the densities on the two sides of the shock line. Making the substitutions $u_n = v\sin(\alpha - \omega)$, $u_t = -v\cos(\alpha - \omega)$ and putting

$$\beta = \frac{v}{u_t} + \frac{4}{(\gamma + 1)G_0} \tag{13}$$

in (10) it now follows from (10), (11) and (12) that

$$\frac{\partial p(\omega, S_*)}{\partial S} = \frac{[(m^2 - 1)\beta^2 \sin^2(\alpha - \omega) - 1]\rho^{-2}G_0}{-\gamma\tau\left(\frac{\rho}{\rho_1} + \frac{\rho_1}{\rho} - 2\right) m^2 \cos(\alpha - \omega)} \tag{14}$$

Equation (14) gives the required formula for the derivative $\partial p/\partial S$ along the shock line. As mentioned before it is now seen that this derivative depends only on the inclination ω of the stream lines (or in consequence of (3) on the inclination α of the shock line). Using the relation

$$\frac{\rho v^2}{p_1} = \gamma M^2 \left(\frac{\rho}{\rho_1}\right) \frac{\cos^2 \alpha}{\cos^2(\alpha - \omega)},$$

which is easily derived, it follows that (1) can be given the form

$$\frac{\theta(\omega)}{p_1} = \frac{p}{p_1} + I \log \Delta \tag{15}$$

where,

$$I = \frac{M^2(\rho/\rho_1)\cos^2\alpha\,[(m^2 - 1)\beta^2 \sin^2(\alpha - \omega) - 1]G_0}{(\gamma - 1)\left[\frac{\rho}{\rho_1} + \frac{\rho_1}{\rho} - 2\right] m^2 \cos^3(\alpha - \omega)}.$$

The left member of (15) is the ratio of the pressure $\theta(\omega)$ along the profile to the pressure p_1 of the incident flow. The ratios p/p_1 and ρ/ρ_1 are given by (5). To evaluate $\log \Delta$ we have the equation (4) and G_0 can be evaluated from the explicit formula for this quantity.[3] Finally the Mach Number m is given by

$$m^2 = \frac{v^2}{c^2} = \frac{v^2}{u_n^2}\cdot\frac{u_n^2}{c^2} = \frac{u_n^2/c^2}{\sin^2(\alpha - \omega)}$$

in which the ratio u_n^2/c^2 has been calculated previously (see, "Calculation of the Curvatures of Attached Shock Waves," *loc. cit.*, ref. 3). Using these formulas in conjunction with the relation (3), it is possible to find, by means of (15), the pressure ratio $\theta(\omega)/p_1$ at any point of the curved profile.

4. Calculations and Conclusions.—The following table shows the values of the quantity $1000I \log \Delta$ at the indicated % chords and for the indicated angles of attack a in the case of the flow past the circular arc profile considered in our previous communication.[1] In carrying out these calculations α was chosen as the independent variable and computations were made for whole and intermediate angular values of this variable over the required ranges. It was found in general that the calculated numerical values of I increased with α but that deviations from this behavior occurred over a

restricted range corresponding to a range of values of ω about $\omega = 0$. Since it appeared that these irregularities were due to the limitation in the accuracy of the trigonometric tables used and since the deviations in question were small, it was decided to replace the calculated values of I in this α range by values obtained on the assumption that I could be considered to vary linearly with α over this range. The values of $1000I \log \Delta$ at the various % chords in the table were then obtained from calculated values of this quantity by interpolation.

% CHORD	$\alpha = -10$	$\alpha = 0$	$\alpha = 26$
0.0	0.00	0.00	0.00
4.0	5.75	0.51	0.87
12.0	11.96	1.20	3.06
21.5	14.64	1.58	5.95
31.0	14.38	1.72	9.80
40.5	13.48	1.65	14.42
50.5	11.38	1.53	19.82
59.0	11.21	1.45	24.57
68.5	10.72	1.37	29.01
78.5	10.00	1.48	31.62
87.5	9.29	1.79	27.42
97.0	8.65	2.48	.....

It is seen that as we go down the columns for each value of α in the table the values of the quantity $I \log \Delta$ at first increase and then decrease. Actually in the case of the column marked $\alpha = 0$ a certain oscillatory behavior in the values of this quantity occurs. However, the quantity $I \log \Delta$, representing the difference between the first approximation of pressure and the higher approximation under consideration, does not seem to have a direct physical significance, and hence it would not appear that this behavior is, in itself, of significance in the physical problem. Since the entries in the table give $1000I \log \Delta$ the values of $I \log \Delta$ are seen to be extremely small and are moreover small in comparison with the values of the first approximation of pressure except near the tail end of the profile for $\alpha = 26$ where they are of the same order of magnitude as the first approximation. It may reasonably be expected that terms involving squares of entropy differences in the gas, which will occur in the next higher approximation, will be small in comparison with the quantity $I \log \Delta$; if this is correct it would appear that the values of pressure over the profile given by the first approximation are very near, except possibly for extreme values of the angle of attack, to the exact theoretical values under the assumption of zero viscosity and thermal conductivity.

[1] "First Approximation of Pressure Distribution on Curved Profiles at Supersonic Speeds," these PROCEEDINGS, **35**, 617–627 (1949).

[2] Prepared under Navy Contract N6onr-180, Task Order V, with Indiana University.

[3] The formulas for the f^1_m, the A_{ij} and the following quantity $G_0(M, \alpha)$ are given in the paper "On Curved Shock Waves," *J. Math. Physics*, **26**, 62-68 (1947). A more explicit formula for G_0 is to be found in the article "Calculation of the Curvatures of Attached Shock Waves," *Ibid.*, **27**, 279-297 (1949).

[4] This relation is derived in §7 of the forthcoming article "The Determination of Pressure on Curved Bodies Behind Shocks," *Communications on Applied Mathematics, New York Univ., Inst. for Math. and Mech.*

[5] See §5 of ref. 4.

---

*THE SELECTIVE ADVANTAGE OF AN ADENINELESS DOUBLE MUTANT OVER ONE OF THE SINGLE MUTANTS INVOLVED**

BY MARY B. MITCHELL AND HERSCHEL K. MITCHELL

KERCKOFF LABORATORIES OF BIOLOGY, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA

Communicated by G. W. Beadle, December 28, 1949

In discussions of evolutionary specialization it is often suggested that if one reaction in a biosynthesis is blocked, the ability to carry out other reactions preceding this one in the series will tend to be lost as a result of mutation, provided that the intermediates involved are not themselves useful. Such a tendency would be predicted if mutations which result in loss of synthetic ability were more frequent, or of greater selective advantage than those which restore the ability. The possibility that loss of synthetic ability may sometimes confer a selective advantage provides an interpretation of the behavior of the adenineless strain of Neurospora which is discussed here. In each of three stock cultures of this strain there occurred spontaneously a second mutation to adenineless, and the cultures became genetically pure for the double mutants. Selection of the double mutants may be accounted for by the fact that, on the medium used, the double mutant in each reaches its maximum rate of growth more quickly than does the original single mutant.

Detection of the Spontaneous Mutations.—A striking characteristic of the strain in which the new mutations appeared makes detection of the presence of certain other adenineless mutant genes quite simple. The accumulation by this strain (isolation number 35203) of a purple pigment is prevented if any one of three other adenineless mutant genes (isolation numbers 27663, 28610 and 44411) is present.[1] Available evidence indicates that the purple pigment is a derivative of an intermediate in the biosynthesis of adenine and that accumulation of the pigment is prevented by the 27663, 28610 and 44411 mutations because these interfere with reactions which come earlier in the series, thereby cutting off the supply of pigment precursor.[1] Hence

the single mutant, 35203, is purple while the double mutant of 35203 with 27663, 28610 or 44411 is not purple.

The three isolates of ad-p (purple-adenineless), which were observed to have lost the purple character but to have retained the requirement for adenine, had been kept as stock cultures at room temperature and had been transferred at six- to eight-week intervals. Some time after the disappearance of the purple character they were out-crossed to wild type and segregation of two adenineless mutants, one of which produced the purple pigment, was demonstrated in all perithecia tested. The double mutant in each case was not purple.

Comparison of Growth Rates of the Single and Double Mutants.—The genetic constitutions of the three double mutants obtained from the above crosses were checked by out-crossing them to wild type. Also double

TABLE 1

STRAIN	GROWTH RATES (MM./HR.) COMPLETE MEDIUM	MINIMAL MEDIUM + ADENINE*
ad-p, isolate 1	2.9	3.5
ad-p, isolate 2	2.6	3.4
ad-p, isolate 3	2.8	3.4
ad-p, isolate 4	2.6	3.4
ad-p, isolate 5	2.8	3.5
ad-x1, ad-p	3.5	3.7
ad-x2, ad-p	3.5	3.6
ad-x3, ad-p	3.5	3.9
27663, ad-p	3.5	3.5
44411, ad-p	3.6	3.8
ad-p, isolate 1 + ad-x1, ad-p	3.5	3.6
ad-p, isolate 1 + ad-x2, ad-p	3.5	3.6
ad-p, isolate 1 + ad-x3, ad-p	3.5	3.8
ad-p, isolate 1 + 27663, ad-p	3.5	
ad-p, isolate 1 + 44411, ad-p	3.5	

* Adenine sulfate, 0.5 mg. per 10 ml. medium.

mutants of ad-p with two of the not-purple mutants, 27663 and 44411, were prepared. Growth rates were measured at 25°C. on complete medium in growth tubes.[3] Innocula were obtained from complete agar slant cultures which were from 7 to 10 days old. The tubes were marked at the mycelial frontier 15 to 20 hours after inoculation and the growth rate determined for the following 48-hour period. Results are given in table 1, where the three spontaneous mutations are designated as ad-x1, ad-x2 and ad-x3. It will be seen that on the complete medium the single mutant ad-p grows more slowly than any of the double mutants. Also, in tubes which were inoculated with a mixture of spores from ad-p and any one of the double mutants the rate of growth was the same as that of the double mutant. No purple pigment was formed in these tubes, except near the

point of inoculation. The growth rate of the purple strain is not constant during this 48-hour period, but is increasing, and after 60 to 70 hours approaches that of the other strains. However, it would appear that, in mixed cultures, the purple strain is forced out very early, since the pigment has been observed only very near the point of inoculation and the rate of growth is fairly constant 15 or 20 hours after inoculation.

The complete medium used had the same composition as that upon which the stock cultures had been kept. It consisted of the usual minimal medium[3] supplemented with an autolyzate of wild-type Neurospora mycelium, prepared as described by Lein, Mitchell and Houlahan.[4] A concentration of autolyzate equivalent to 40 mg. moist weight of mycelium per ml. of medium was used. On minimal medium supplemented with adenine there is little difference in the growth rate of ad-p and that of the other strains in the 48-hour period considered, but there is a greater difference during the first 15 hours after inoculation, the distance along the tube covered by ad-p being about two-thirds of that covered by the other strains. Apparently the difference is sufficient to allow selection of the double mutant on this medium too, since no pigment was observed in mixed cultures except near the innoculum.

In liquid culture, if the autolyzate is used as a supplement, after 30 to 40 hours growth the dry weight of mycelium from the double mutants and mixed cultures is 4 to 5 times greater than that from the purple mutant. After 50 to 60 hours the difference in dry weight is only about 16%, but no pigment was observed in the mixed cultures. When adenine was used as a supplement (1 mg. adenine sulfate per 20 ml. of medium) there was little difference in the dry weights after 40 hours, and the mixed cultures were purple.

Genetic Relationships of the Spontaneous Mutants and the Induced, Not-Purple Adenineless Mutants.—From crosses of the double mutants to wild type the three new mutants, ad-x1, -x2 and -x3 were obtained without ad-p. They were crossed with each other, with 27663, 28610 and 44411, and with a fourth not-purple adenineless mutant, 3254, which gives a faintly purple double mutant with ad-p. The crosses were made on corn meal agar slants and the ascospores obtained were suspended in sterile water and spread on agar minimal medium in petri-plates. The plates were kept at 25°C. for 12 to 14 hours after heat treatment and then examined for wild-type spores. These can be distinguished from mutant spores since the latter produce very short haphae and then stop growing on the unsupplemented medium, while growth of the wild type continues.[4]

From the results of these crosses, given in table 2, it appears that the three spontaneous mutants are all genetically different and that ad-x2 is also different from the four induced not-purple mutants. It is possible that ad-x1 is allelic with 28610 and ad-x3 with 27663; fertile crosses were not obtained in these two cases.

Discussion.—It is clear that in mixed cultures, containing ad-p and a double mutant of ad-p with one of the not-purple adenineless mutants described, the double mutant is selected on either of the two agar media used. The slower initial growth rate characteristic of ad-p appears to be adequate reason for the selection, but is possible that other less obvious factors are involved. The cause of the reduced growth rate is not known. Possibly the accumulation of pigment is in some way responsible but this has not been demonstrated. It can only be said that a mutation which prevents the accumulation of pigment removes, or compensates for, a detrimental effect of the ad-p mutation.

Fries, working with Ophiostoma, has observed three cases in which a mutant strain, after having spontaneously acquired a second mutation, has become genetically pure for the double mutant. Two of these strains were guanine-requiring and the second mutation in both was to hypozanthineless.[5] The third was an adenineless mutant in which a mutation to biotinless occurred.[6] These, of course, differ from the case reported here in that new growth requirements are introduced, but knowledge of possible interrelations in the biosyntheses concerned might diminish this difference.

TABLE 2

STRAINS	WILD-TYPE PROGENY OBSERVED	STRAINS	WILD-TYPE PROGENY OBSERVED
ad-x1 × ad-x2	++	ad-x2 × 27663	++
ad-x1 × ad-x3	++	ad-x2 × 28610	++
ad-x1 × 3254	++	ad-x2 × 44411	++
ad-x1 × 27663	++	ad-x3 × 3254	++
ad-x1 × 28610	No fertile cross	ad-x3 × 27663	No fertile cross
ad-x1 × 44411	++	ad-x3 × 28610	++
ad-x2 × ad-x3	++	ad-x3 × 44411	++
ad-x2 × 3254	++		

One may question the usefulness of speculation upon the applicability of selection of a further loss of synthetic ability as a mechanism for evolutionary specialization. The observations of Fries and those described here serve to establish that under certain laboratory conditions the phenomenon of selection does take place. Possibly it does so to a significant extent under natural circumstances as well.

It is of interest to compare the mutations which prevent the accumulation of purple pigment by ad-p with suppressors which operate on Drosophila mutants characterized by abnormally high pigment production. If the adenineless strains of Neurospora were always kept on a complex medium, and the adenine requirement not therefore recognized, then the not-purple mutations which prevent pigment production would, if they were encountered, be looked upon as suppressors of the purple character, quite analogous to suppressors of black,[7] sable[8] and perhaps others, in Drosophila. There

might be found then, at least four complete suppressors and one partial suppressor of purple, namely 27663, 28610, 44411, ad-x2 and 3254. Furthermore, there is another genetically distinct adenineless mutant, 44206, previously described,[1] but not otherwise discussed, here, which produces the purple pigment under different conditions and to a lesser extent. Hence this strain is quite different in appearance from ad-p and it might therefore be classified as an unrelated mutant suppressed by the same series of suppressors which operate on ad-p.

Summary.—In three stock cultures of an adenineless strain of Neurospora a double mutant arose as the result of spontaneous mutation to adenineless, and, in each case, the double mutant forced out the original single mutant. Under conditions which existed at the time the selection took place the double mutants show a faster initial rate of growth than the original single mutant. The three new mutations are genetically distinct and each one, in combination with the original mutant prevents accumulation of a purple pigment which is characteristic of this strain. Cultures from mixed innocula of the purple mutant with any one of the double mutants have the growth rate of the double mutant and form no pigment.

Acknowledgment.—The authors are indebted to Mrs. Mary Emerson who first observed that the cultures of ad-p which had lost the purple character had retained the requirement for adenine.

[*] Work supported by grants from the Rockefeller Foundation.

[1] Mitchell, H. K., and Houlahan, M. B., *Fed. Proc.*, **3**, 370 (1946).

[2] Ryan, F. J., Beadle, G. W., and Tatum, E. L., *Am. J. Bot.*, **30**, 784 (1943).

[3] Beadle, G. W., and Tatum, E. L., *Ibid.*, **32**, 678 (1945).

[4] Lein, J., Mitchell, H. K., and Houlahan, M. B., PROC. NATL. ACAD. SCI., **34**, 435 (1948).

[5] Fries, Nils, *Physiologia Plantarum*, **2**, 78 (1949).

[6] Fries, Nils, *Proc. 8th Intern. Congr. Genetics*, Stockholm (in press), 1949.

[7] Plough, H. H., *Verh. d. V. Intern. Kongr. f. Vererbung, Berlin I.*, **1**, 193 (1928).

[8] Schultz, J., and Bridges, C. B., *Am. Nat.*, **66**, 323 (1932).

HIGHER PROPERTIES OF PHYSICAL SYSTEMS OF CURVES

BY EDWARD KASNER AND JOHN DE CICCO

DEPARTMENTS OF MATHEMATICS, COLUMBIA UNIVERSITY, NEW YORK, AND ILLINOIS INSTITUTE OF TECHNOLOGY AND DE PAUL UNIVERSITY, CHICAGO

Communicated December 10, 1949

1. A *system* S_∞ of ∞^2 curves in a field of force consists of curves along which a constrained motion is possible such that the pressure P is proportional to the normal component N of the force vector. Thus $P = kN$ where $k(\neq -1)$ is the constant factor of proportionality.[1]

. A system S_k, where $k \neq -1$ but may be infinite, of ∞^3 curves is defined by the ordinary differential equation of third order

$$(\psi - y'\phi)y''' = [\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2]y'' - \left[3\phi - \frac{2k(\phi + y'\psi)}{(1 + k)(1 + y'^2)}\right]y''^2, \quad (1)$$

where $\phi(x, y)$ and $\psi(x, y)$ are the horizontal and vertical components of the force vector.

The important special cases of physical interest are (*a*) dynamical trajectories, $k = 0$; (*b*) general catenaries, $k = 1$; (*c*) generalized brachistochrones, $k = -2$; (*d*) velocity curves, $k = \infty$.

2. At any lineal element (x, y, y'), the speed of a particle traversing those curves of a system S_k that have four-point contact with their circle of curvature is called the *hyperosculating speed*.

The hyperosculating speed v at any lineal element (x, y, y') is determined by the formula

$$v^2 = \frac{(3 + k)(\psi - y'\phi)(\phi + \psi y')}{\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2}. \quad (2)$$

The corresponding hyperosculating radius of curvature is

$$r = \frac{(3 + k)(\phi + \psi y')(1 + y'^2)^{1/2}}{(1 + k)[\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2]}. \quad (3)$$

3. Consider a lineal elément E through a given point P, which is not perpendicular to the direction of the line of force through P. In a system S_k where $k \neq -1, -3$, there is a *single* curve, actual or virtual, which passes through E such that it has four-point contact with its circle of curvature at E.

PROPERTY III. *The locus of the centers of the ∞^1 hyperosculating circles constructed at a given point P for all elements is a conic passing through P in the direction of the force vector.*

The equation of this conic in running coordinates (X, Y) is

$$(1 + k)[\phi_y(X - x)^2 + (\psi_y - \phi_x)(X - x)(Y - y) - \psi_x(Y - y)^2] + (3 + k)[-\psi(X - x) + \phi(Y - y)] = 0. \quad (4)$$

PROPERTY IV. *The normal to the conic (4) of Property III at the point P cuts the conic again at a distance equal to $(3 + k)/(1 + k)$ times the radius of curvature of the line of force which passes through P.*

These results were originally stated by Kasner in his Princeton Colloquium Lectures. In the present paper, we shall discuss some new results which are related to the Properties III and IV.

4. At any point P, the *exceptional directions* of a given field of force are given by the quadratic equation in y',

$$\psi_x + (\psi_y - \phi_x)y' - \phi_y y'^2 = 0. \tag{5}$$

A point P is *exceptional* if every direction through P is exceptional. At a non-exceptional point P, there are two exceptional directions, real or imaginary.

In a field of force, which is neither constant nor elastic, there is a net of $2\infty^1$ exceptional curves. This exceptional net is orthogonal if and only if the field of force is conservative.

The direction of the line of force at a point P is exceptional if and only if its curvature is zero at P. The direction orthogonal to that of the line of force at P is exceptional if and only if the magnitude of the force vector at P is instantaneously stationary upon moving P in this orthogonal direction.

For a field of force, the exceptional net consists of its lines of force and their orthogonal trajectories if and only if the lines of force are straight and the magnitude of the force vector is constant along each of the orthogonal trajectories.

All central fields of force belong to the above class.

The hyperosculating speed v for any system S_k where $k \neq -1, -3, \infty$, is a finite non-zero constant if and only if the lines of force are straight lines orthogonal to a fixed proper, or null, or rectilinear circle Γ such that the magnitude of the force vector at any point P is inversely proportional to the distance PQ where Q is the inverse or anti-inverse point of P with respect to Γ.

5. A point P is *ordinary* if it is non-exceptional and if the hyperosculating speed v at P depends not only on P but also on the direction y' through P.

Through a non-exceptional point P, consider a direction which is neither exceptional nor the direction of the line of force nor its orthogonal direction. If v_1 and v_2 denote the hyperosculating speeds in this direction corresponding to the two systems S_{k1} and S_{k_2}, respectively, then

$$\frac{v_1^2}{v_2^2} = \frac{3 + k_1}{3 + k_2}. \tag{6}$$

6. As the direction y' varies about an ordinary point P, the corresponding hyperosculating speed v assumes various values. The directions, along which this speed v is instantaneously stationary, satisfy the quadratic equation in y',

$$[\psi_x(\psi^2 - \phi^2) - \phi\psi(\psi_y - \phi_x)] + 2\phi\psi(\phi_y - \psi_x)y' + [\phi_y(\psi^2 - \phi^2) - \phi\psi(\psi_y - \phi_x)]y'^2 = 0. \tag{7}$$

These are called the *principal directions*. At a non-ordinary point P, every direction is a principal direction. At an ordinary point P, there are only two principal directions, which may be real or imaginary.

For a field of force, which is not of the types described in Section 4, there is a principal net of $2\infty^1$ curves.

At an ordinary point P, the hyperosculating speed v is instantaneously stationary along each of the two principal directions (7), *which are symmetrically situated with respect to the line of force.*

The following formula is analogous to Euler's formula for the normal curvatures of a surface.

At an ordinary point P, let λ_j, *for* $j = 1, 2$, *denote the angle between any direction* y' *and a principal direction* y_j, *and let* v *and* v_j *denote the hyperosculating speeds corresponding to these directions* y' *and* y_j'. *Then*

$$\frac{v^2 - v_1^2}{v_1^2 \sin^2 \lambda_1} = \frac{v^2 - v_2^2}{v_2^2 \sin^2 \lambda_2}. \tag{8}$$

Our final result is related to the Property IV stated above.

At a direction of a non-exceptional point P, which is neither exceptional nor the direction orthogonal to that of the line of force, the ratio ρ *of the hyperosculating curvature* κ *of a system* S_k *to that of the hyperosculating curvature* κ_0 *of the velocity system* S_∞ *is*

$$\rho = \frac{\kappa}{\kappa_0} = \frac{1 + k}{3 + k}. \tag{9}$$

Let the direction in the above result be taken as that of the line of force. Then κ_0 is the curvature of the line of force and κ is the curvature of the rest trajectory of the system S_k. Thus this result reduces to the theorem of Kasner concerning the comparative curvatures of the rest trajectory of a system S_k and the line of force.[2]

[1] Kasner, "Differential Geometric Aspects of Dynamics," Princeton Colloquium Lectures, Published by *Am. Math. Soc.*, 1913, 1934, 1948. See also a series of papers by Kasner and De Cicco published in these PROCEEDINGS, 1947–1949. The five characteristic properties of trajectories and general physical systems are there discussed. Property I relates to the circular focal locus, Property II to the angular relation, Property V to the variation of the conic of Property III, Properties III and IV are stated in the present paper. Projective characteristics have been given by Terracini for trajectories.

[2] Kasner and Mittleman, "A General Theorem on the Initial Curvatures of Dynamical Trajectories," these PROCEEDINGS, **28**, 48–52 (1942). Kasner and De Cicco, "Physical Curves in Generalized Fields of Force, *Ibid.*, **33**, 169–172 (1948). De Cicco, "Constrained Motion upon a Surface under a Generalized Field of Force," *Bull. Am. Math. Soc.*, **53**, 993–1001 (1947).

THEORETICAL RELATIONSHIPS AMONG SOME MEASURES OF CONDITIONING

BY CONRAD G. MUELLER

COLUMBIA UNIVERSITY

Communicated by C. H. Graham, December 10, 1949

The relationships among the various measures of strength of conditioning constitute an important problem for conditioning theory. Many different measures have been used.[1] The measures *latency* and *magnitude* are based on the occurrence of a single response, while *number of responses in extinction* and *the rate of responding* in a "free-response" situation are based on more than one instance of a response. *Probability of response occurrence* is another term that is encountered in the literature; it is used most frequently in cases where more than one response is possible (e.g., right and left turns in a T maze) and in circumstances when it is possible to compute the frequency or the percentage of times that a specified response is given. Percentage of response occurrence is taken to be an estimate of the probability of obtaining the response.

Some theoretical formulations are concerned with one or two measures of strength; others are more inclusive. In only few cases has an attempt been made to present a theory of the relation among measures. In most treatments that consider several measures, the relations among the measures are empirically determined.

The purpose of the present note is to indicate one possible theoretical account of the relationships among latency of response, rate of responding and the probability of occurrence of a response. The last measure serves as the starting point for the discussion and provides the terms in which the other concepts are related.

Consider the Skinner bar-pressing situation[2] in which a rat's responses may occur at any time and at any rate during the period in which the animal is in the experimental cage. Assume that the responses under constant testing conditions are randomly distributed in time. Let the rate of occurrence of these responses be represented by r. It may then be shown that the probability, $P_{>t}$, of obtaining an interval between two responses greater than t is

$$P_{>t} = e^{-rt} \tag{1}$$

where e is the base of Naperian logarithms.[3] The probability of obtaining n responses in an interval, T, is

$$P_n = (rT)^n e^{-rT}/n! \tag{2}$$

Equation (1) gives us a statement of the distribution of time intervals associated with various rates of responding. For example, for the median

time interval $P_{>t}$ is 0.5 and $-rt$ is $\log_e 0.5$ or the median t is $0.69/r$. Equation (2) gives the probability of various numbers of responses within some specified time interval. For example, the probability of getting exactly one response in an interval, T, is $(rT)e^{-rT}$. The relation between equations (1) and (2) is obvious when we consider the probability of getting no responses in an interval, T. In this case P_0 is e^{-rT}.

Equations (1) and (2) permit us to transform a rate measure into a probability measure. Since we are dealing with a continuous distribution (time), the probability of a response at any particular time is zero, but the probability of a response within given time intervals is finite and specifiable.

Latency usually refers to the time interval between a stimulus and a response and thus is not directly considered in the previous development. Assume, however, that the stimulus conditions are one determinant of the rate of responding, that is, that the rate has different values for different stimulus conditions. This assumption is consistent with the discussions by Skinner and others who have emphasized the measurement of rate; the assumption would presumably be an elementary requirement for any measure.

Under the circumstances of the assumption, t may be employed in discussing latency, since the latter would be the time interval between the beginning of the observation period (when a stimulus was presented) and the first response. Thus, on the assumption that stimulus conditions are a determinant of rate of responding and on the previous assumption that the responses are randomly distributed in time, a statement of the rate of responding under specified stimulus conditions implies a probability statement of the delay of length t between the presentation of the stimulus and the occurrence of the first response. This statement tells us not only of the distribution of latencies but also of the relationship between some representative value, say the median latency, and the rate of responding; for example, the probability of a response greater than the median latency, t_{md}, is 0.5; and from equation (1) we see that $-rt_{md} = \log_e 0.5$ or that the median latency equals $0.69/r$.

The preceding development does not imply any particular theory of conditioning but may be incorporated into a large class of theories. For example, if the foregoing discussion is combined with a theory that states that rate of responding is proportional to the number of responses that remain to be given in extinction, the measure of number of responses in extinction is immediately related to our latency and probability terms. In other words, if

$$r = k(N-n), \qquad (3)$$

where N is the number of responses in extinction, n is the number of re-

sponses already given, and k is a constant, we may substitute $k(N-n)$ for r in equation (1) and obtain

$$P_{>t} = e^{-k(N-n)t}. \qquad (4)$$

This equation may then be examined for relationships existing among the terms n, N, P and t. In addition to the relationships among latency, rate and number of responses in extinction, equation (4) may be used to predict the distribution of responses in extinction for a constant strength and the distribution of time intervals between responses at various stages of extinction.

Since the present argument follows mainly from the assumption of a random distribution of responses in time, it is of interest to examine data

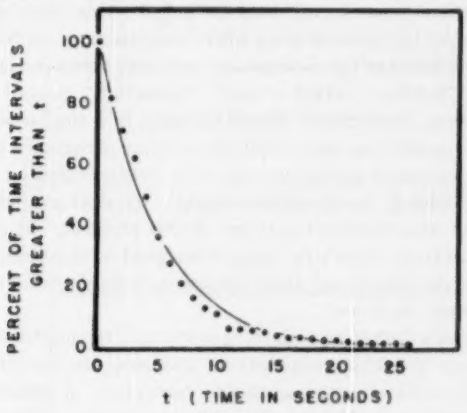


FIGURE 1

The percentage of inter-response time intervals greater than t, where t is time in seconds. The data are from an experiment with white rats in a bar-press situation as described in the text. The line drawn through the data is a plot of equation (1).

for direct evidence of randomness as well as for evidence relating to the above outlined consequences of randomness.

The data in figure 1 were taken from measurements obtained during the course of periodic reconditioning.[4] The data represent the responses of a single animal during a 20-minute session of "three-minute" periodic reconditioning. Within this observation period the rate of responding was approximately constant. The question at issue is whether the responses in this interval are distributed randomly. Equation (1) states that the probability of getting an interval between responses greater than t is e^{-rt}, where r is the rate of responding expressed in the same units as t. In the

20-minute session, 238 responses were made, 237 time intervals were recorded, and the rate in this session is 0.20 response per second. Thus, without direct reference to the distribution of time intervals, theory specifies the distribution of time intervals between responses uniquely. In this case the probability of getting a time interval greater than t (in seconds) is $e^{-0.20t}$. The ordinate of figure 1 shows the percentage of the intervals between responses that were greater than the various time values specified on the abscissa. The solid line through the date in figure 1 represents the theoretical function. The data are consistent with the assumption that the measured responses occurred randomly in time.

Although the data of figure 1 may be representative of the agreement between data and theory under the conditions specified, certain cases of systematic deviations from theory may be noted. One class of deviations, for example, may be found in cases where animals show marked "holding" behavior, i.e., where the bar is depressed and held down for many seconds. Although the "holding" period is not a "refractory" period[4] in the usual sense of the term, it obviously affects the data in a similar way. During the "holding" period, the probability of response occurrence is zero. One complicating feature in analyzing responses characterized by "holding" is the fact that "holding" is of variable length. The data available at present do not warrant an extensive treatment of this problem, but the simplicity that may result from apparatus changes designed to eliminate the factor of "holding" and the advantages that may accrue from the additional response specification may be shown.

An example of a distribution showing systematic deviations from theory is shown in figure 2. The computations and plot are similar to those in figure 1. The ordinate represents the percentage of intervals between responses greater than the specified abscissa values. The solid line is theoretical. The constant of the line was determined, as in the case in figure 1, directly from the rate of responding without reference to the distribution of time intervals. The fit is obviously poor; the function appears sigmoid and asymmetric.

Let us assume that the analysis leading to equation (1) and applied to figure 1 is correct when applied to all portions of the observation period except the time spent in "holding." An additional test may then be applied to the data from which figure 2 was obtained. Now we are interested in the measurement of the time interval between the end of one response and the beginning of the next.[5] Figure 3 shows the results of such measurements in the form of a plot of the percentage of intervals between the end of one response and the beginning of the next that were greater than the specified abscissa values. The solid line through the data is theoretical when the rate term, r, is set equal to the ratio of the number of responses to the total time minus the "holding" time, i.e., to the number of responses

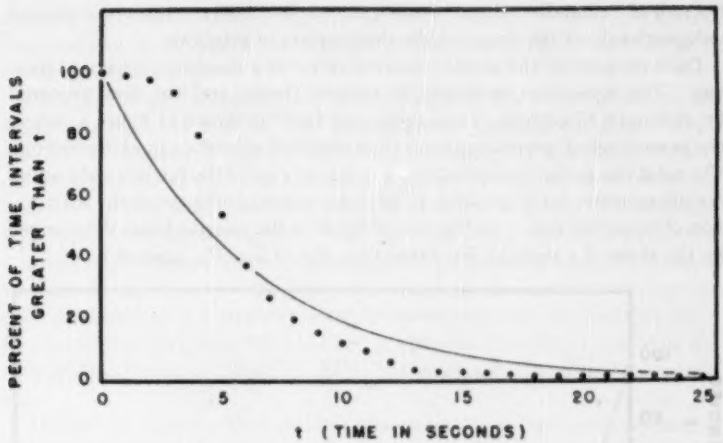


FIGURE 2

A plot similar to figure 1 showing the deviation from theory in cases of "holding" behavior. The line drawn through the data is a plot of equation (1).

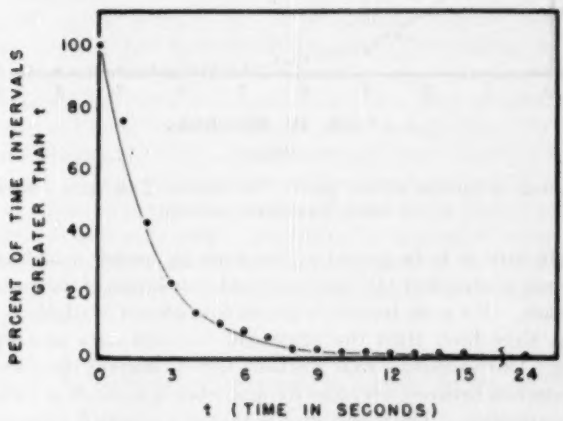


FIGURE 3

The data of figure 2 "corrected for holding." The plot is similar to that in figures 1 and 2, except that the measured interval is the time between the end of one response and the beginning of the next response. The line drawn through the data is a theoretical one described in the text.

per unit of "available" time. As in figures 1 and 2 the constant is evaluated independently of the shape of the distribution of intervals.

Data relevant to the present analysis of latency measures are not numerous. The agreement between the present theory and the data reported by Felsinger, Gladstone, Yamaguchi and Hull[6] is shown in figure 4, where the percentage of latencies greater than specified abscissa values are plotted. The solid line is the theoretical curve. In the case of the latency data under consideration it is not possible to evaluate r independently of the distribution of time intervals. In the case of figure 4 the constant was determined by the slope of a straight line fitted to a plot of $\log_e P_{>t}$ against t.

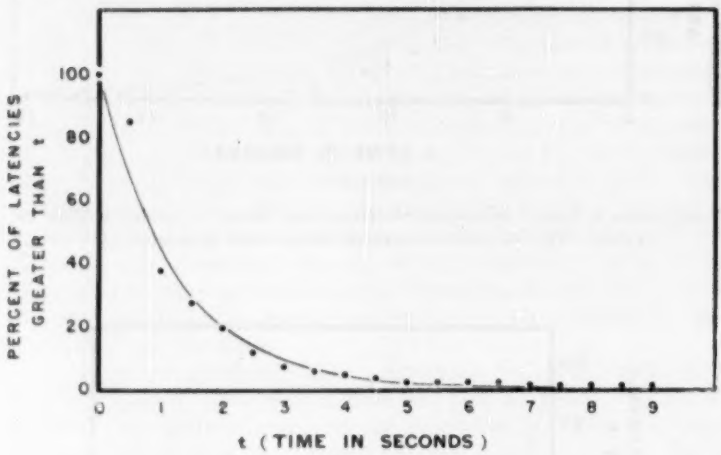


FIGURE 4

The percentage of latencies greater than t. The data are from figure 1 of Felsinger, Gladstone, Yamaguchi and Hull.[6]

Probably little is to be gained at this time by further sampling of the consequences of equation (1), but many additional tests of the formulation may be made. For some tests appropriate data are not available. For the tests that have been tried the agreement between data and theory is promising. One prediction that has been tested concerns the distribution of time intervals between responses for a number of animals at comparable stages in extinction. The expectation is that at a specified stage in extinction the intervals between, say, response R_n and R_{n+1}, for a large number of animals, will be distributed in a manner similar to that shown in figure 1 and that the constant, r (therefore the steepness of drop of the curve) will vary systematically with n. In other words, the steepness of the drop of a curve such as found in figure 1 will depend on where in extinction the inter-

vals are measured. In fact this expectation seems to be borne out by the cases measured, although the number of measurements at each stage of extinction is not large.

Finally, it may be pointed out that the form of the present account has important consequences for the treatment of experimental data. Since one of the features of the account is the possibility of specifying the frequency distributions of the measures discussed it is possible to eliminate many of the problems associated with the arbitrary selection of representative values in summarizing data. On the basis of the preceding equations, one may state changes in one statistic, say the arithmetic mean, in terms of changes in another, say the geometric mean or the median. Therefore, data using different statistics are made comparable and the multiplicity of functions that may arise from the use of different descriptive statistics not only ceases to pose a difficult problem but is actually an aid to theory testing.

Summary.—A theoretical account of some relationships among measures of strength of conditioning has been considered. (1) If we assume that responses in a "free-response" situation are randomly distributed in time, we obtain directly a statement of the probability of occurrence of a response (or of any number of responses) within a specified time interval as a function of the length of the interval and of the rate of responding; we also obtain a statement of the probability of occurrence of inter-response time intervals of varying lengths. (2) If we assume that, for any specified stimulus condition, there corresponds some rate of responding, it turns out that the probability of occurrence of latencies of various lengths may be specified for various rates of responding, or, for a fixed probability value, the relation between latency and rate may be specified. (3) Finally, where these considerations are added to a theory specifying the relationship between rate of responding and number of responses yet to occur, the number of responses in extinction may be related to the latency and probability terms as well as to rate. In addition to statements about average values, the present formulation has consequences for the distribution of time intervals between responses and, by extension, for the distribution of latency measures.

[1] Hull, C. L., *Principles of Behavior*, D. Appleton-Century Co., New York, 1943.

[2] Skinner, B. F., *The Behavior of Organisms*, D. Appleton-Century Co., New York, 1938.

[4] A slightly different equation results if we assume that a "refractory" period exists, i.e., that immediately after a response there is a period during which the probability of getting a response is zero. If we assume that the transition from the "refractory" period to randomness is instantaneous, the probability of getting an interval greater than t is

$$P_{>t} = e^{-r(t-t_0)}$$

where t_0 is the "refractory" period. The formulation is more complex if the transition is treated as a gradual one or if the "refractory" period has a variable length.

[4] The data reported here were recorded by Mr. Michael Kaplan in the Psychological Laboratories of Columbia University.

[5] This is merely a first approximation. Subsequent analyses may show that the interval between the end of one response and the beginning of the next is not independent of the "holding" period. The results of our procedure indicate that the approximation is useful for the present.

[6] The experiment by Felsinger, Gladstone, Yamaguchi and Hull [*J. Exptl. Psychol.*, **37**, 214–228 (1947)] may not provide an optimal test of our formulation for two reasons. The first is that the data are reported in a frequency distribution with step intervals which begin at zero. If the shortest latency were greater than zero, starting the step intervals at the lowest measure would be more appropriate. The use of zero as a lower limit could easily make an exponential distribution more normal. The method of summarizing the data may account for the deviation of the point at 0.5 second in figure 4. The deviation of this point is an expression of the fact that the distribution reported by Felsinger, Gladstone, Yamaguchi and Hull does not have a maximum frequency at the first step interval.

In the second place, it may be assumed that the many transient discriminative stimuli associated with the exposure of the bar may play a more important rôle than the continuous ones associated with the presence of the bar. Although it is possible to extend the present notion to stimuli of short duration which end before the occurrence of the response, additional assumptions are required. A less equivocal test of the present theory may be expected from a distribution of latencies obtained from an experimental procedure of the sort used by Skinner (op. cit.), Frick [*J. Psychol.*, **26**, 96–123 (1948)] and others. After a period of, say, no light, a light is presented and stays on until one response occurs (Skinner) or stays on for some fixed period of time sufficiently long to insure the occurrence of many responses (Frick). Such experimental procedures would minimize unspecified transient stimuli and would parallel more closely the notion that stimulus conditions determine a rate of responding. The procedure used by Frick has the additional advantage of permitting the measurement of the time interval between the onset of the stimulus and the first response and the subsequent intervals between responses under "the same" stimulus conditions.

PARTIAL DIFFERENTIAL EQUATIONS AND GENERALIZED ANALYTIC FUNCTIONS

By Lipman Bers

Syracuse University and The Institute for Advanced Study

Communicated by Marston Morse, December 24, 1949

1. Introduction.—In this note (an abstract of results to be published in full elsewhere) we present the fundamentals of a theory of complex-valued functions $f = u(x, y) + iv(x, y)$ whose real and imaginary parts are connected by the equations

$$u_x = \sigma(x, y)v_y, \; u_y = -\sigma(x, y)v_x, \tag{1}$$

a theory which parallels closely that of analytic functions.

As early as 1891 Picard[1] foresaw the possibility of such a theory, but his suggestion seems to have remained unnoticed. Several years ago Gelbart and the author[2] considered the case $\sigma = a(x)b(y)$. In this case there exists an integration process (used implicitly by Beltrami[3] in a special case) which yields explicit formulas for solutions of (1) corresponding to the powers of a complex variable. From these all other single-valued regular solutions can be constructed, at least locally, by means of "formal power series." Our methods could be extended to systems of higher order (Diaz[4]) and to equations in three-space (Protter[5]), but they seem to be inadequate for the treatment of singular and multiple-valued solutions. The general case (1) has been considered recently by Bergman[6] and by Markushevitch (as reported by Petrovskii[7]), and especially important results have been obtained by Položii.[8] In the present paper we do not construct solutions in closed form as was done in the special case mentioned above, but are able to obtain considerably stronger general results. We shall impose only very mild regularity conditions on the coefficient $\sigma(x, y)$, but in order to achieve as close an analogy as possible with classical function theory we shall require at present that these conditions be satisfied in the whole function-theoretical plane. Most of our results, however, imply theorems of a local character.

It is hardly necessary to point out that the analogy between analytic functions and solutions of partial differential equations of elliptic type has inspired many recent investigations, notably the work of Bergman and Schiffer[9] on the kernel-function, and the theory of mappings by means of solutions of systems of differential equations developed by Lavrentyeff, Schapiro, and Gergen and Dressel.[10]

2. *Pseudo-Analytic Function.*—We set $x + iy = z$ and write functions of x and y as functions of z, without implying by this an analytic dependence on z. A function $\sigma(x, y) = \sigma(z)$ defined in a bounded domain $\boldsymbol{D}$ is called admissible if $\sigma > 0$, and σ_x and σ_y exist and satisfy a Hölder condition. A function σ defined over a Riemann surface $\boldsymbol{F}$ is admissible, if in the neighborhood of each point of $\boldsymbol{F}$ it is admissible as a function of the local parameter. In the following we use a fixed function σ defined and admissible over the whole complex plane (Riemann sphere). Thus $\sigma(z)$ is admissible for $|z| < +\infty$ and $\rho(\zeta) = \sigma(1/\zeta)$ is admissible for $|\zeta| < +\infty$.

We say that a complex-valued function $f(z) = u(x, y) + iv(x, y)$ is pseudo-analytic with respect to σ at a point z_0 if in some neighborhood of this point the partial derivatives of u and v exist, are continuous and satisfy (1). A function is pseudo-analytic in a domain if it is pseudo-analytic at every point in the domain. Complex constants are pseudo-analytic, so are linear combinations of pseudo-analytic functions with real coefficients.

Since equations (1) preserve their form under a conformal transformation, pseudo-analyticity can be defined over an arbitrary Riemann surface.

We shall make almost no use of this fact at present, except for the case of the Riemann sphere.

With each function $g(z) = \varphi(x, y) + i\psi(x, y)$ we associate the "reduced" function ${}^{\sigma(z)}g(z) = \sigma(x, y)^{-1/2}\varphi(x, y) + i\sigma(x, y)^{1/2}\psi(x, y)$. Using this notation we may express pseudo-analyticity by a differentiability requirement.

THEOREM. *A function $f(z)$ defined in a bounded domain $\boldsymbol{D}$ is pseudo-analytic in $\boldsymbol{D}$ if and only if for every z_0 in $\boldsymbol{D}$ the function ${}^{\sigma(z)}[f(z) - f(z_0)]$ (considered as a function of the complex variable z) is differentiable at z_0.*

The proof of this theorem requires much of the theory developed below.

3. Behavior at a Point.—The following theorem contains and sharpens some of the recent results by Položii. Our proof uses a method due to Carleman.[11]

THEOREM. *Let $f(z)$ be single-valued, pseudo-analytic and not a constant for $0 < |z - z_0| < r$. Then either $f(z)$ is pseudo-analytic at z_0 and ${}^{\sigma(z)}[f(z) - f(z_0)] \sim a(z - z_0)^n$, where $a \neq 0$ and n is a positive integer, or for some positive integer n, ${}^{\sigma(z)}f(z) \sim a(z - z_0)^{-n}$, $a \neq 0$, or $f(z)$ comes arbitrarily close to any given value in every neighborhood of z_0.*

In the first case we say that $f(z)$ assumes the value $f(z_0)$ with multiplicity n, in the second that $f(z)$ has at z_0 a pole of order n. In the third case (essential singularity) Picard's theorem holds by virtue of a result by Grötzsch,[12] since mappings by pseudo-analytic functions are mappings of bounded eccentricity.

COROLLARY. *Pseudo-analytic functions are interior transformations in the sense of Stoïloff.*

The behavior of a pseudo-analytic function at $z = \infty$ may be described in a similar manner.

4. Formal Powers.—Let r be a real rational number, $r \neq 0$, $|r| = p/q$, where p and q are relatively prime positive integers. Let a and ζ be complex numbers, $a \neq 0$. We say that a function $f(z)$ is a formal power with exponent r, coefficient a and center at ζ, and write $f(z) = Z^{(r)}(a, \zeta; z)$, if $f(z)$ is q-valued and pseudo-analytic for $0 < |\zeta - z| < +\infty$, if $w = f(z)$ is a one-to-one mapping of the q-times covered z-plane (with branch-points at ζ and ∞) onto the p-times covered w-plane (with branch-points at 0 and ∞) such that $f(\zeta) = 0, f(\infty) = \infty$ if $r > 0, f(\zeta) = \infty, f(\infty) = 0$ if $r < 0$, and if for $z \to \zeta$ we have that ${}^{\sigma(z)}f(z) \sim a(z - \zeta)^r$. We also set $Z^{(0)}(a, \zeta; z) = a$, $Z^{(r)}(0, \zeta; z) = 0$.

THEOREM. *For every rational r and complex a and ζ the formal power $Z^{(r)}(a, \zeta; z)$ exists and is uniquely determined.*

It is easy to see that $Z^{(r)}(a + b, \zeta; z) = Z^{(r)}(a, \zeta; z) + Z^{(r)}(b, \zeta; z)$, and that for every real α, $Z^{(r)}(\alpha a, \zeta; z) = \alpha Z^{(r)}(a, \zeta; z)$.

With the aid of the formal powers we may associate with every function $f(z)$ which is pseudo-analytic at ζ the "differential quotients"

$$D_\zeta^0 f = f(\zeta),\ D_\zeta^n f = n! \lim_{z \to \zeta} \frac{\left\{ f(z) - \sum_{\nu=0}^{n-1} Z^{(\nu)}(D_\zeta^\nu f/\nu!, \zeta; z) \right\}}{(z-\zeta)^n}, \qquad n = 1, 2, \ldots.$$

These differential quotients always exist, though $f(z)$ considered as a function of x and y need not have partial derivatives of order higher than the second.

THEOREM. *If $f(z)$ is pseudo-analytic at ζ and $D_\zeta^n f = 0$, $n = 0, 1, \ldots$, then $f(z) \equiv 0$.*

It is easy to see that the uniform limit f of a sequence of pseudo-analytic functions $\{f_j\}$ is pseudo-analytic. It can be shown that if $f_j \to f$ uniformly in a domain $\boldsymbol{D}$, then $D_\zeta^\nu f_j \to D_\zeta^\nu f$, the convergence being uniform with respect to ζ in every compact subset of $\boldsymbol{D}$.

5. Rational Functions.—We say that a pseudo-analytic function is a rational function if it has no singularities except poles, an entire rational function if it has no singularities except perhaps a pole at infinity.

THEOREM. *Every non-constant rational function has as many zeros as it has poles (provided poles and zeros are counted with proper multiplicities). Every rational function is a sum of formal powers with integral exponents. If $\zeta_i (i = 0, 1, \ldots, R)$ are distinct points, $n_i (i = 1, 2, \ldots, R)$ non-vanishing integers such that $n_1 + n_2 + \ldots + n_R = 0$ and b a non-vanishing complex number, then there exists one and only one rational function assuming at ζ_0 the value b, having at $\zeta_i (i = 1, 2, \ldots, R)$ a zero of order n_i if $n_i > 0$ and a pole of order $-n_i$ if $n_i < 0$, and having no other zeros or poles.*

If $f(z)$ is an entire rational function and ζ a given complex number, then $f(z)$ may be written as a "formal polynomial with center ζ":

$$f(z) = \sum_{\nu=0}^{n} Z^{(\nu)}(a_\nu, \zeta; z). \tag{2}$$

We say that this formal polynomial is of degree n if $a_n \neq 0$. It then has exactly n zeros.

THEOREM. *A formal polynomial of degree n is uniquely determined by its center, its leading coefficient and its n zeros. These parameters may be prescribed arbitrarily.*

Instead of prescribing the zeros of (2) we may require that $f(z)$ should satisfy the conditions $D_\zeta^\nu f = b_{\nu j}, \nu = 0, 1, \ldots, N_j, j = 1, 2, \ldots, R$, provided that $\zeta_j \neq \zeta_l$ for $j \neq l$, and $N_1 + N_2 + \ldots + N_R + R = n$.

6. Formal Power Series and the Expansion Theorem.—A formal power series is a series of the form

$$\sum_{\nu=0}^{\infty} Z^{(\nu)}(a_\nu, \zeta; z). \tag{3}$$

Together with (3) we consider the ordinary power series $\sum a_\nu z^\nu$ and denote its radius of convergence by r.

THEOREM. *The series (3) converges absolutely and represents a pseudo-analytic function for $|z - \zeta| < r$, converges uniformly for $|z - \zeta| \leq r' < r$ diverges for $|z - \zeta| > r$.*

A similar theorem holds for series of the form $\sum Z^{(-\nu)}$.

THEOREM. *Let $f(z)$ be pseudo-analytic for $|z - \zeta| < R$ and set $a_\nu = D_\zeta^\nu f/\nu!$. There exists a positive number α depending only on the function σ, such that the series (3) converges and represents the function $f(z)$ for $|z - \zeta| < \alpha R$.*

Thus if f is an entire function ($R = +\infty$), the series converges for all finite values of z. A similar expansion theorem holds for functions pseudo-analytic for $|z - \zeta| > R$.

7. *Cauchy's Formula.*—The second order equations resulting from (1) are

$$(u_x/\sigma)_x + (u_y/\sigma)_y = 0, \qquad (\sigma v_x)_x + (\sigma v_y)_y = 0. \tag{4}$$

Using the so-called fundamental solutions of these equations the values of u and v at an interior point of a domain may be expressed by line-integrals extended over the boundary of the domain and involving the boundary values of u and v (Bergman, Položii).

We are also able, however, without using explicitly the fundamental solutions, to obtain a Cauchy type formula formally identical with the classical one.

We note first that $Z^{(\nu)}(a, \zeta; z)$ depends continuously on a and ζ (at least for $\zeta \neq z$). Hence if C is a rectifiable arc admitting the parametric representation $\zeta = \zeta(s)$, $|\zeta'(s)| = 1$, $s_0 \leq s \leq s_1$, and $A(\zeta)$ a continuous function defined on C, the integral

$$\int_C Z^{(\nu)}[A(\zeta)d\zeta, \zeta; z] = \int_{s_0}^{s_1} Z^{(\nu)}\{A[\zeta(s)]\zeta'(s), \zeta; z\}\, ds$$

is defined for every z not on C.

THEOREM. *Let $\boldsymbol{D}$ be a domain bounded by a simple closed rectifiable curve C, $f(z)$ a function which is pseudo-analytic in $\boldsymbol{D}$ and continuous on C. Then*

$$\frac{1}{2\pi i}\int_C Z^{(-1)}\{[\sigma(\zeta)f(\zeta)]id\zeta, \zeta; z\} = f(z)$$

if z in an interior point of $\boldsymbol{D}$. The integral vanishes if z is an exterior point of $\boldsymbol{D}$.

We proceed to enumerate some consequences of this theorem.

8. *Isolated Singularities.*—THEOREM. *Let $f(z)$ be single-valued and pseudo-analytic for $0 < |\zeta - z| < R$. Then*

$$f(z) = \sum_{\nu=-\infty}^{+\infty} Z^{(\nu)}(a_\nu, \zeta; z),$$

the expansion being convergent for $0 < |\zeta - z| < \alpha R$.

Here α is the constant of the second theorem in §6. The number of non-vanishing a_{-p} determines in the usual way the character of the point ζ.

9. Approximation Theorem.—Let $\boldsymbol{D}$ be a bounded domain bounded by the closed rectifiable curves $C_0, C_1, \ldots C_N$, C_0 being the outer boundary curve. Let ζ_0 be a point exterior to C_0 (which may be the point at infinity), ζ_j ($j = 1, 2, \ldots, N$) a point interior to C_j. In $\boldsymbol{D}$ we consider a single-valued pseudo-analytic function $f(z)$.

THEOREM. *There exists a sequence of rational pseudo-analytic functions having no poles except at the points $\zeta_0, \ldots, \zeta_N$, which converges to $f(z)$ in $\boldsymbol{D}$, the convergence being uniform in every compact subset of $\boldsymbol{D}$.*

COROLLARY. *If $\boldsymbol{D}$ is simply connected, f may be expanded in a series of formal polynomials.*

Recently Eichler[13] established the following theorem for certain linear partial differential equations of elliptic type in n-space with analytic coefficients. Let $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, $\boldsymbol{D}_3$ be bounded domains homeomorphic to the interior of the n-sphere, such that $\overline{\boldsymbol{D}}_1 \subset \boldsymbol{D}_1$, $\overline{\boldsymbol{D}}_2 \subset \boldsymbol{D}_2$, $\overline{\boldsymbol{D}}$ denoting the closure of $\boldsymbol{D}$. Let φ be a solution of the equation considered defined in $\boldsymbol{D}_2$ and ϵ a positive number. Then there exists a solution ψ defined in $\boldsymbol{D}_3$ such that $|\varphi - \psi| < \epsilon$ in $\boldsymbol{D}_1$. A more special theorem of this kind has been obtained previously by Bergman.[14] Our result shows that for equations of the form (4) Eichler's theorem holds without analyticity assumptions.

10. The Logarithmic Function.—Let ζ' and ζ'' be two distinct points, and set

$$L(a, \zeta', \zeta''; z) = \int_{\zeta}^{\zeta''} Z^{(-1)}\{[^{1/\sigma(\zeta)}a]d\zeta, \zeta; z\},$$

the integration being performed along some fixed path.

THEOREM. *The function $f(z) = L(a, \zeta', \zeta''; z)$ is pseudo-analytic for $z \neq \zeta', \zeta''$ and is such that the difference $f(z) - [a \log (z - \zeta') - a \log (z - \zeta'')]$ is single-valued and $O(|\log|(z - z')(z - z'')||)$. Any other function with these properties differs from $f(z)$ by a constant.*

From this theorem it is not difficult to infer the existence of a function $f(z) = L(a, \zeta, \infty; z)$ which is pseudo-analytic for $z \neq \zeta, \infty$ and differs from $a \log (z - \zeta)$ by a single-valued function $= O(|\log|z - z||)$. If a is real, then $u = \boldsymbol{Re}\, L(a, \zeta', \zeta''; z)$ is a fundamental solution of the first equation (4) for every domain containing ζ' but not ζ''.

11. Multiple-Valued Functions.—THEOREM. *Let $f(z)$ be a k-valued pseudo-analytic function defined for $0 < |\zeta - z| < r$. Then*

$$f(z) = \sum_{p=-\infty}^{+\infty} Z^{(p/k)}(a_p, \zeta; z),$$

the expansion being convergent at all points sufficiently near to and distinct from ζ.

If only a finite number of the coefficients $a_{-1}, a_{-2}, \ldots$, are different from

zero, we say that ζ is an algebraic singularity of $f(z)$. Algebraic singularities at $z = \infty$ are defined similarly. A function with only algebraic singularities is called algebraic.

THEOREM. *The sum of all determinations of an algebraic pseudo-analytic function is a pseudo-analytic rational function.*

If $f(z)$ is algebraic, there is associated with it a Riemann surface on which it is single-valued. The genus of this (closed) Riemann surface will be called the genus of $f(z)$.

THEOREM. *Let $f(z)$ be an algebraic pseudo-analytic function of genus p. Then it admits a parametric representation of the form*

$$z = \varphi(t), f = \omega(t)$$

where $\varphi(t)$ is a single-valued analytic function and $\omega(t)$ a single-valued function which is pseudo-analytic with respect to an admissible function $\rho(t)$. If $p = 0$, then $\rho(t)$ is admissible over the whole t-plane, and φ and ω are rational. If $p = 1$, ρ is admissible for $|t| < +\infty$ and all three functions ρ, φ, ω are doubly-periodic with the same periods. If $p > 1$, then ρ is admissible for $|t| < 1$, and all three functions ρ, φ, ω are automorphic with respect to the same group of hyperbolic motions.

While this theorem is almost self-evident (as is the corresponding uniformization theorem for non-algebraic functions) we state it in order to show how the investigation of multiple-valued solutions of (1) leads naturally to the theory of pseudo-analytic functions on Riemann surfaces. This theory should also include equations of the form (1) with multiple-valued $\sigma(x, y)$ and the case where $\sigma(x, y)$ becomes discontinuous or vanishes at certain points and on certain lines.

[1] Picard, E., *C.R. (Paris)*, **112**, 168–188 (1891).

[2] Bers, L., and Gelbart, A., *Quart. Appl. Math.*, **1**, 168–188 (1943); *Trans. Am. Math. Soc.*, **56**, 67–93 (1944); *Ann. Math.*, **48**, 342–357 (1947).

[3] Beltrami, E., *Opere mat.*, vol. 3, Milano, 1911, pp. 115–128, 349–377.

[4] Diaz, J. B., *Am. J. Math.*, **68**, 611–659 (1946).

[5] Protter, M. H., *Trans. Am. Math. Soc.*, **63**, 314–349 (1948).

[6] Bergman, S., *Ibid.*, **61**, 452–498 (1947); *Duke Math. J.*, **14**, 349–366 (1947).

[7] Petrovskii, I. G., *Uspekhi Mat. Nauk*, **1**, 44–70 (1946).

[8] Položii, G. N., *C.R. (Doklady)*, **58**, 1275–1278 (1947); **60**, 769–772 (1948); *Mat. Sbornik N.S.*, **24** (**66**), 375–384 (1949).

[9] Bergman, S., and Schiffer, M., *Bull. Am. Math. Soc.*, **53**, 1141–1151 (1947); *Duke Math. J.*, **15**, 535–566 (1948).

[10] Lavrentyeff, M. A., *Mat. Sbornik*, **42**, 407–423 (1935); *N.S.*, **21** (**63**), 513–554 (1947); *Izvestya Akad. Nauk SSSR, Ser. Mat.*, **12**, 513–554 (1948); Schapiro, Z., *C.R. (Doklady)*, **30**, 690–692 (1941). Gergen, T. T., and Dressel, F. G., *Bull. Am. Math. Soc.*, **54** 1062, (1948). See also a paper by J. J. Gergen and F. G. Dressel to appear in the *Duke Math. J.*

[11] Carleman, T., *C.R. (Paris)*, **197**, 471–474 (1933).

[12] Grötzsch, H., *Leipziger Ber.*, **80**, 503–510 (1928).

[13] Eichler, M., *Math. Z.*, **49**, 565–575 (1944).

[14] Bergman, S., *Duke Math. J.*, **6**, 537–561 (1940).

THE EXCEPTIONAL SIMPLE LIE ALGEBRAS F_4 AND E_6

BY CLAUDE CHEVALLEY AND R. D. SCHAFER

DEPARTMENTS OF MATHEMATICS, COLUMBIA UNIVERSITY AND UNIVERSITY OF PENNSYLVANIA

Communicated by A. A. Albert, December 28, 1949

Let K be an algebraically closed field of characteristic 0. The exceptional simple Jordan algebra $\mathfrak{J}$ over K is the (non-associative) algebra of dimension 27 whose elements are 3×3 Hermitian matrices

$$X = \begin{pmatrix} \xi_1 & x_3 & \bar{x}_2 \\ \bar{x}_3 & \xi_2 & x_1 \\ x_2 & \bar{x}_1 & \xi_3 \end{pmatrix} \tag{1}$$

with elements in the Cayley algebra $\mathfrak{C}$ of dimension 8 over K, multiplication being defined by

$$X \circ Y = {}^1/_2(XY + YX)$$

where XY is the ordinary matrix product. We write the trace

$$\xi_1 + \xi_2 + \xi_3 = Sp\, X.$$

The derivation algebra $\mathfrak{D}$ of $\mathfrak{J}$ is the Lie algebra of endomorphisms D of $\mathfrak{J}$ satisfying

$$D(X \circ Y) = DX \circ Y + X \circ DY. \tag{2}$$

By a right multiplication R_Y is meant the endomorphism

$$X \to X \circ Y \qquad \text{for every } X \text{ in } \mathfrak{J}.$$

We characterize the exceptional simple Lie algebras of dimension 52 and 78 over K by means of $\mathfrak{J}$ as follows:

THEOREM. *The exceptional simple Lie algebra F_4 of dimension 52 and rank 4 over K is the derivation algebra $\mathfrak{D}$ of the exceptional Jordan algebra $\mathfrak{J}$ of dimension 27 over K. The exceptional simple Lie algebra E_6 of dimension 78 and rank 6 over K is the Lie algebra*

$$\mathfrak{D} + \{R_Y\}, \qquad Sp\, Y = 0, \tag{3}$$

spanned by the derivations of $\mathfrak{J}$ and the right multiplications of elements Y of trace 0.

We begin by proving a Principle of Triality for the elements of the orthogonal Lie algebra $\mathfrak{o}(8, K)$ of all 8×8 skew-symmetric matrices over K. We regard the elements of $\mathfrak{o}(8, K)$ as endomorphisms of the Cayley algebra $\mathfrak{C}$. This algebra is equipped with a trace function $Sp\, x = x + \bar{x}$ satisfying

$$Sp\ xy = Sp\ yx, \qquad Sp\ x(yz) = Sp\ (xy)z\ (= Sp\ xyz).$$

Also $Sp\ xy$ is a non-degenerate bilinear form. The alternative law in $\mathfrak{C}$ gives the formula

$$Sp\ (sx)(yz) + Sp\ (ys)(zx) = Sp\ (xy)(sz + zs) \text{ for } x, y, z, s \text{ in } \mathfrak{C}. \quad (4)$$

For, writing $[x, y, z] = x(yz) - (xy)z$, we have $[x, y, z] = [z, x, y]$, so that $x(yz) + (zx)y = z(xy) + (xy)z$ and $Sp\ (sx)(yz) + Sp\ (ys)(zx) = Sp\ s(x(yz)) + Sp\ ((zx)y)s = Sp\ (x(yz) + (zx)y)s = Sp\ (z(xy) + (xy)z)s = Sp\ (xy)(sz + zs)$.

An endomorphism U of $\mathfrak{C}$ is in $\mathfrak{o}(8, K)$ if and only if U leaves the norm form $x\bar{x}$ invariant:

$$(Ux)\bar{x} + x(\overline{Ux}) = 0. \quad (5)$$

The derivations D of $\mathfrak{C}$ satisfy (5); so do the left and right multiplications $L_s: x \to sx$ and $R_t: x \to xt$ of elements s and t of trace 0. It follows that every element U in $\mathfrak{o}(8, K)$ may be written uniquely in the form

$$U = D + L_s + R_t, \qquad \bar{s} = -s, \bar{t} = -t. \quad (6)$$

For if $D + L_s + R_t = 0$, then $t = -s$ and $L_s - R_s$ is a derivation of $\mathfrak{C}$: $s(yz) - (ys)z = (sy)z - (ys)z + y(sz) - y(zs)$ for all y, z in $\mathfrak{C}$. That is, $[s, y, z] + [y, z, s] - [y, s, z] = 3[s, y, z] = 0$, and s associates with all y, z in $\mathfrak{C}$. Then $s = 0$, $D = L_s = R_t = 0$. Since $\mathfrak{o}(8, K)$ is of dimension 28, while the derivation algebra of $\mathfrak{C}$ is the exceptional simple Lie algebra G_2 of dimension 14, every element of $\mathfrak{o}(8, K)$ is a sum (6).

PRINCIPLE OF TRIALITY. *For U in $\mathfrak{o}(8, K)$, there exist unique U', U'' in $\mathfrak{o}(8, K)$ such that*

$$Sp\ (Ux)yz + Sp\ x(U'y)z + Sp\ xy(U''z) = 0 \quad (7)$$

for all x, y, z in $\mathfrak{C}$.

For $U = D$, take $U' = U'' = D$. Then $0 = Sp\ D(xyz) = Sp\ (Dx)yz + Sp\ x(Dy)z + Sp\ xy(Dz)$. For $U = L_s$, take $U' = R_s$, $U'' = -L_s - R_s$. Then $Sp\ (sx)(yz) + Sp\ x(ys)z - Sp\ (xy)(sz + zs) = Sp\ (sx)(yz) + Sp\ (ys)(zx) - Sp\ (xy)(sz + zs) = 0$ by (4). Similarly for $U = R_t$, take $U' = L_t$, $U'' = -L_t - R_t$. But any U in $\mathfrak{o}(8, K)$ may be written in the form (6), so the existence of U', U'' satisfying (7) is assured. To show the uniqueness, we show that, if $Sp\ x(U'y)z + Sp\ xy(U''z) = 0$ identically, we have $U' = U'' = 0$. Since $Sp\ xy$ is a non-degenerate bilinear form, $Sp\ x((U'y)z + y(U''z)) = 0$ for all x, y, z implies

$$(U'y)z + y(U''z) = 0 \qquad \text{for all } y, z \text{ in } \mathfrak{C}. \quad (8)$$

Let $z = 1$ in (8); then $U'y = yu$ for u in $\mathfrak{C}$ with $Sp\ u = 0$. Let $y = 1$ in (8): then $U''z = vz$ for v in $\mathfrak{C}$. Hence (8) becomes

$$(yu)z + y(vz) = 0 \qquad \text{for all } y, z \text{ in } \mathfrak{C}. \tag{9}$$

Putting $y = z = 1$ in (9), we have $v = -u$, $[y, u, z] = 0$, or u associates with all y, z in C. Then $u = 0$, $U' = U'' = 0$.

Associated with the exceptional simple Jordan algebra $\mathfrak{J}$ are the bilinear form $Sp\, X \circ Y$ and the trilinear form

$$\phi(X, Y, Z) = Sp\,(X \circ Y) \circ Z = Sp\, X \circ (Y \circ Z).$$

Any derivation D of $\mathfrak{J}$ leaves these forms invariant:

$$Sp\, DX \circ Y + Sp\, X \circ DY = 0, \tag{10}$$

$$\phi(DX, Y, Z) + \phi(X, DY, Z) + \phi(X, Y, DZ) = 0. \tag{11}$$

Conversely, any endomorphism D of $\mathfrak{J}$ leaving both $Sp\, X \circ Y$ and $\phi(X, Y, Z)$ invariant is a derivation of $\mathfrak{J}$. For (10) implies $Sp\,(D(X \circ Y)) \circ Z + \phi(X, Y, DZ) = 0$. With (11) this gives

$$Sp\,(DX \circ Y + X \circ DY - D(X \circ Y)) \circ Z = 0 \tag{12}$$

identically. Since $Sp\, X \circ Y$ is a non-degenerate bilinear form, (12) implies (2), D is a derivation of $\mathfrak{J}$.

Denote by e_i the matrix (1) with $\xi_i = 1$, all remaining entries zero. Let $\mathfrak{T}_i$ be the set of matrices (1) with all entries except x_i and $\bar{x}_i$ zero; $\mathfrak{T}_i$ is then in a natural one-to-one correspondence with $\mathfrak{C}$. If $T_i \in \mathfrak{T}_i$, then $e_i \circ T_i = 0$, $e_j \circ T_i = \frac{1}{2}T_i$ if $j \neq i$. We propose first to determine all derivations which map e_1, e_2 and e_3 upon 0. The spaces $\mathfrak{T}_i$ are invariant under such a derivation D, since $0 = D(e_i \circ T_i) = (De_i) \circ T_i + e_i \circ (DT_i) = e_i \circ (DT_i)$ implies DT_i is in $\mathfrak{T}_i + Ke_j + Ke_k$ where $\{i, j, k\} = \{1, 2, 3\}$, while $\frac{1}{2}DT_i = D(e_j \circ T_i) = (De_j) \circ T_i + e_j \circ (DT_i) = e_j \circ (DT_i)$ implies DT_i is in $\mathfrak{T}_i + \mathfrak{T}_k$. Let U be the restriction of D to $\mathfrak{T}_1$. Since D leaves invariant the quadratic form

$$Sp\, X^2 = \sum \xi_i^2 + 2\sum x_i\bar{x}_i,$$

we see that U leaves invariant the norm form $x_1\bar{x}_1$, so U is in $\mathfrak{o}(8, K)$. Also the restrictions U' and U'' of D to $\mathfrak{T}_2$ and $\mathfrak{T}_3$ are in $\mathfrak{o}(8, K)$. But since D leaves invariant the cubic form

$$Sp\, X^3 = \sum \xi_i^3 + 2\sum \xi_i(x_j\bar{x}_j + x_k\bar{x}_k) + 3Sp\, x_1x_2x_3,$$

it follows that U, U', U'' satisfy (7). Hence any derivation of $\mathfrak{J}$ which maps e_1, e_2, e_3 upon 0 has the form

$$X \to \begin{pmatrix} 0 & U''x_3 & U'\bar{x}_2 \\ U''\bar{x}_3 & 0 & Ux_1 \\ U'x_2 & U\bar{x}_1 & 0 \end{pmatrix} \tag{13}$$

for X in (1), where U is in $\mathfrak{o}(8, K)$ and U', U'' are the corresponding elements of $\mathfrak{o}(8, K)$ given by the Principle of Triality. Conversely, any mapping (13) is a derivation of $\mathfrak{J}$ since it leaves invariant both $Sp\ X \circ Y$ and $\phi(X, Y, Z)$. The mappings (13) then give 28 linearly independent derivations of $\mathfrak{J}$, and actually form a subalgebra of $\mathfrak{D}$ isomorphic to $\mathfrak{o}(8, K)$.

The space of all elements of trace 0 of $\mathfrak{J}$ which are mapped upon 0 by R_{e_i} is the space $\mathfrak{T}_i' = \mathfrak{T}_i + K(e_j - e_k)$. Let $\mathfrak{D}_i$ be the space of derivations of $\mathfrak{J}$ which map e_i upon 0; then the elements of $\mathfrak{D}_i$ map $\mathfrak{T}_i'$ into itself and leave invariant the restriction to $\mathfrak{T}_i' \times \mathfrak{T}_i'$ of the bilinear form $Sp\ X \circ Y$. On the other hand, the only operation of $\mathfrak{D}_i$ which maps $\mathfrak{T}_i'$ upon $\{0\}$ is 0; for such an operation maps e_i, $e_j - e_k$, and $e_i + e_j + e_k$ upon 0, which shows that it belongs to $\mathfrak{o}(8, K)$, and therefore that it is 0 since it maps $\mathfrak{T}_i$ upon 0. The algebra $\mathfrak{D}_i$ is therefore isomorphic with a subalgebra of $\mathfrak{o}(9, K)$. On the other hand, if X and Y are in $\mathfrak{J}$, then

$$[R_X, R_Y] = R_X R_Y - R_Y R_X$$

is a derivation, because $\mathfrak{J}$ is a Jordan algebra.[3] In particular, if $T_1 \in \mathfrak{T}_1$, the operation $[R_{T_1}, R_{e_2 - e_3}]$ is in $\mathfrak{D}_1$ and maps $e_2 - e_3$ upon T_1. Since the operations of $\mathfrak{o}(8, K)$ map $e_2 - e_3$ upon 0, we have dim $\mathfrak{D}_1 \geq 28 + 8 = 36 =$ dim $\mathfrak{o}(9, K)$, and $\mathfrak{D}_1$ is isomorphic with $\mathfrak{o}(9, K)$; the same is of course true of $\mathfrak{D}_2$ and $\mathfrak{D}_3$.

Let D be any derivation of $\mathfrak{J}$. Then $De_1 = De_1^2 = 2e_1 \circ De_1$, whence $De_1 = T_2 + T_3 \in \mathfrak{T}_2 + \mathfrak{T}_3$. We can find operations D_2 and D_3 of $\mathfrak{D}_2$ and $\mathfrak{D}_3$, respectively, such that $D_2 e_1 = T_2$, $D_3 e_1 = T_3$, and D_2 and D_3 are then uniquely determined modulo $\mathfrak{o}(8, K)$. The operation $D - (D_2 + D_3)$ is then in $\mathfrak{D}_1$; this proves that the algebra $\mathfrak{D}$ of derivations of $\mathfrak{J}$ is $\mathfrak{D}_1 + \mathfrak{D}_2 + \mathfrak{D}_3$ and that its dimension is $(36 - 28) + (36 - 28) + 36 = 52$. The adjoint representation of $\mathfrak{D}$ induces a representation of $\mathfrak{o}(8, K)$ which is the sum of the adjoint representation and of the representations given by the spaces $\mathfrak{D}_i/\mathfrak{o}(8, K)$ $(i = 1, 2, 3)$. If we map every $D \in \mathfrak{D}_i$ upon $D(e_j - e_k)$ (where (i, j, k) is an even permutation of $(1, 2, 3)$), then we obtain equivalences of the representations given by the spaces $\mathfrak{D}_1/\mathfrak{o}(8, K)$, $\mathfrak{D}_2/\mathfrak{o}(8, K)$, $\mathfrak{D}_3/\mathfrak{o}(8, K)$ with the representations $U \to U$, $U \to U'$, and $U \to U''$ constructed above, and it is easily seen that these representations are inequivalent to each other and to the adjoint representation of $\mathfrak{o}(8, K)$. Thus the space $\mathfrak{D}$, considered as a representation space of $\mathfrak{o}(8, K)$, splits into four inequivalent simple representation spaces. It follows that any ideal $\mathfrak{a}$ in $\mathfrak{D}$ must be the sum of 0, 1, 2, 3 or 4 of these spaces. On the other hand, each $\mathfrak{D}_i$ being simple, the intersection of an ideal with $\mathfrak{D}_i$ is either $\{0\}$ or $\mathfrak{D}_i$. It follows immediately that the only ideals in $\mathfrak{D}$ are $\{0\}$ and $\mathfrak{D}$; that is, $\mathfrak{D}$ is simple. Since there is only one simple Lie algebra of dimension 52, namely F_4, we see that $\mathfrak{D}$ is the algebra F_4. The set $\mathfrak{J}_0$ of elements of trace 0 in $\mathfrak{J}$ is mapped into itself by the operations of $\mathfrak{D}$ and yields a representation of de-

gree 26 of $\mathfrak{D}$. Considered as a representation space of $\mathfrak{o}(8, K)$, $\mathfrak{J}_0$ splits into two invariant spaces of dimension 1 (spanned by $e_1 - e_2$ and $e_1 - e_3$, respectively) and three inequivalent spaces of dimension 8 (namely, $\mathfrak{T}_1$, $\mathfrak{T}_2$ and $\mathfrak{T}_3$). Any invariant subspace of $\mathfrak{J}_0$ with respect to $\mathfrak{D}$ of dimension > 2 must therefore contain one of the spaces $\mathfrak{T}_i$. If it contains $\mathfrak{T}_1$, then, being invariant with respect to $\mathfrak{D}_1$, it contains $e_2 - e_3$, and, being invariant with respect to $\mathfrak{D}_2$ and $\mathfrak{D}_3$, it contains $\mathfrak{T}_2$ and $\mathfrak{T}_3$. It follows immediately that $\mathfrak{J}_0$ gives a simple representation of $\mathfrak{D}$.

If X and Y are in $\mathfrak{J}$, then $[R_X, R_Y]$ is in $\mathfrak{D}$; moreover, if $D \in \mathfrak{D}$, then $[D, R_Y] = R_{DY}$ and $Sp\, DY = 0$. It follows that (3) is a Lie algebra $\mathfrak{E}$. Since $D(1) = 0$ for every derivation D of $\mathfrak{J}$, $D + R_Y = 0$ implies $Y = 0$, $D = 0$; thus $\mathfrak{E}$ is of dimension 78, and the adjoint representation of $\mathfrak{E}$ induces a representation of $\mathfrak{D}$ which is the sum of the adjoint representation and of the representation whose space is the set $\mathfrak{R}$ of right multiplications of elements in $\mathfrak{J}_0$. Since $\mathfrak{D}$ and $\mathfrak{R}$ yield simple representation spaces of $\mathfrak{D}$ of distinct dimensions, the only possible ideals in $\mathfrak{E}$ are $\{0\}$, $\mathfrak{D}$, $\mathfrak{R}$ and $\mathfrak{D} + \mathfrak{R} = \mathfrak{E}$; but $\mathfrak{D}$ and $\mathfrak{R}$ are obviously not ideals, which proves that $\mathfrak{E}$ is a simple Lie algebra. The space $\mathfrak{J}$ yields an obviously simple representation of degree 27 of $\mathfrak{E}$. Now, outside E_6, the only simple Lie algebras of dimension 78 are the symplectic algebra $\mathfrak{sp}(12, K)$ and the orthogonal algebra $\mathfrak{o}(13, K)$. Using the formulas of H. Weyl for the computation of the degrees of the simple representations,[2] we find that the degree of the simple representation of $\mathfrak{sp}(12, K)$ of smallest degree > 12 is 65, while the degree of the simple representation of degree > 13 of $\mathfrak{o}(13, K)$ is 64. It follows that $\mathfrak{E}$ is the exceptional simple algebra E_6.

[1] Albert, A. A., "A Structure Theory for Jordan Algebras," *Ann. Math.*, **48**, 550, formula (8) (1947).

[2] Weyl, H., *The Classical Groups*, Chap. VII, §§ 8 and 9. Actually Weyl gives here the explicit formula only for the degrees of the representations of the symplectic group, but the formula for the orthogonal group can be obtained in exactly the same manner.

A TOPOLOGIZED FUNDAMENTAL GROUP[1]

BY J. DUGUNDJI

UNIVERSITY OF SOUTHERN CALIFORNIA

Communicated by Hassler Whitney, December 1, 1949

Introduction.—For a $C^\circ LC^\circ$ Hausdorff space, it is well known that there is a 1-1 correspondence between its covering spaces and the subgroups of its fundamental group.[2] This note arose from an investigation to determine whether non-trivial[3] covering spaces for $C^\circ LC^\circ$ spaces can be constructed, and whether they can be classified in some fashion. Our principal tool is the topologized fundamental group of Hurewicz.[4]

1. The Group χ.—Instead of continuous deformations, we use step-homotopies. Let Y be a topological space and σ a covering of Y by open sets. The path[5] α_1 is σ-homotopic to the path α_n, $\alpha_1\sigma \simeq \alpha_n$, if there exists a finite sequence $\alpha_1, \ldots, \alpha_n$ of paths, all with the same beginning and ending points, such that for every t, $0 \leq t \leq 1$ and each $i = 1, 2, \ldots, n-1$, $\alpha_i(t)$ and $\alpha_{i+1}(t)$ are contained in a common set of σ. Two paths are equivalent if they are σ-homotopic for every σ. With the usual definition of product and inverse of paths, the set of all equivalence classes of paths closed at a point $y_0 \in Y$ forms a group. Denote the equivalence class of α by α. $\chi(Y, y_0)$ is this group, with the topology: given a covering σ of Y, the *σ-nbd* of α, $N(\alpha, \sigma) = \{\beta | \beta\sigma \simeq \alpha\}$; the sets $N(\alpha, \sigma)$ are taken as a basis.[6]

1. $\chi(Y, y_0)$ is a topological group which (a) is zero dimensional, (b) has arbitrary small open normal subgroups, (c) is metrizable, and (d) is completeable.[7] If Y is C°, the collection of groups $\{\chi(Y, y_0)\}$ is a system of local groups.[8]

The position of χ relative to the Poincaré group π_1 and the Vietoris fundamental group[9] V_1 is given by

2. If Y is C° and $\chi = \pi_1$, then Y is[10] ω-LC_1. If Y is a $C^\circ LC^\circ$ ω-LC_1 space with barycentric refinements,[11] then $\chi = \pi_1$. In a compact $C^\circ LC^\circ$ metric space, V_1 is homeomorphic with the completion of χ.

An $F: X \to Y$ yields a continuous homomorphism $F_+ : \chi(X, x_0) \to \chi(Y, F(x_0))$ induced by the correspondence $\alpha : F\alpha$; F_+ is a homotopy invariant.

2. Open Subgroups and Coverings.—Let Y be C° and $\mathfrak{K}$ any open subgroup of $\chi(Y, y_0)$; choose a covering ϑ so that $N(1, \vartheta) \subset \mathfrak{K}$. Two paths α, β starting at y_0 and ending at any $y \in Y$ are equivalent mod $\mathfrak{K}$ if $\alpha\beta^{-1} \in \mathfrak{K}$; the equivalence class of α is written $[\alpha]$. These equivalence classes, for all $y \in Y$ are the points of $\tilde{Y}$. Given α and any open $U \supset \alpha(1)$, U contained in some set of ϑ, define $(\alpha, U) = \{[\alpha\beta] | \beta \subset U\}$; the sets (α, U) are taken as a basis in $\tilde{Y}$. The projection mapping $\omega : \tilde{Y} \to Y$ is defined by $\omega[\alpha] = \alpha(1)$. $\tilde{Y}$ is a Hausdorff space, and

3. Let Y be C°, $\tilde{Y}$ constructed as above. Then $\tilde{Y}$ is C°, ω is continuous, and ω induces a continuous 1-1 mapping of[12] $\tilde{Y}^{-1}(\tilde{y})$ onto $Y^{-1}(\omega\tilde{y})$ for each $\tilde{y} \in \tilde{Y}$. Closed paths at $\tilde{y}_0$ are sent on all those, and on only those closed paths at $\omega\tilde{y}_0 = y_0$ representing elements of $\mathfrak{K}$.

Adding an LC° assumption, $\tilde{Y}$ is a covering space, $\chi(\tilde{Y}, \tilde{y}_0) = \mathfrak{K}$ and we have

THEOREM. *If Y is $C^\circ LC^\circ$ Hausdorff, its covering spaces are in a 1-1 correspondence with the open subgroups of χ. Precisely: (a) if $\tilde{Y}$ is any covering space of Y, then $\chi(\tilde{Y})$ is an open subgroup of $\chi(Y)$; (b) if $\mathfrak{K}$ is any open subgroup of $\chi(Y)$, there exists a covering space with $\chi(\tilde{Y}) = \mathfrak{K}$, and ($c$) any two coverings with homeomorphic χ's are themselves homeomorphic.*

In general then, universal covering spaces[13] for $C^\circ LC^\circ$ spaces do not exist.

3. Closed Subgroups.—With a closed subgroup $B \subset \chi(Y)$ construct a $\hat{Y}$ as follows: the elements and the projection map are defined as in 2. Topology: for any covering σ, and open $U \in \sigma$, define $(\alpha, U, \sigma) = \{[\beta] | \beta(1) \in U, \text{ and } \exists \text{ a } \delta \subset U \text{ with } \beta\sigma \simeq \alpha\delta\}$. The sets (α, U, σ) are taken as a basis. $\hat{Y}$ is a Hausdorff space, and

4. If Y is C°, $\hat{Y}$ as above, then every fiber $\omega^{-1}(y)$ is homeomorphic to the coset space χ/B.

It follows that for any closed B, χ/B is always totally arcwise disconnected.

In addition we have the analog of 3.

5. Let Y be C° and $\hat{Y}$ as above. Then $\hat{Y}$ is C°, ω is continuous, and ω induces a continuous 1-1 mapping of $\hat{Y}^I(\hat{y})$ onto $Y^I(\omega\hat{y})$ for every $\hat{y} \in Y$. Closed paths at $\hat{y}_0$ are sent on all those and on only those closed paths at $\omega\hat{y}_0 = y_0$ representing elements of B. If Y is $C^\circ LC^\circ$, this correspondence is a homeomorphism.

Even with the LC° assumption, $\hat{Y}$ is in general not even a fiber space;[14] however a covering homotopy theorem does hold in the LC° case. The calculation of $\chi(\hat{Y})$ is not known, except that in general, $\chi(\hat{Y}) \neq B$. This also shows that in Y^I, when Y is $C^\circ LC^\circ$, the set of closed paths at $y_0 \in Y$ equivalent to 1 is in general *not* a connected set, in contrast with the result that in Y^I the set of closed paths at $y_0 \in Y$ homotopic to 1 is arcwise connected.

[1] This represents work done mainly under contract N7 onr-384, Office of Naval Research.

[2] See, e.g., Hopf, H., *Math. Ann.*, **102**, 562–623 (1930). For the symbols C^n, LC^n, Kuratowski, C., *Fund. Math.*, **24**, 269–287 (1934).

[3] Y is always a covering space of itself.

[4] Hurewicz, W., *Fund. Math.*, **25**, 467–485 (1935).

[5] A path in Y is a continuous map of $0 \leq t \leq 1$ into Y.

[6] If Y is compact metric, χ is the Hurewicz group, *loc. cit.*

[7] In the sense of van Dantzig, D., *Math. Ann.*, **107**, 587–626 (1932).

[8] Steenrod, N., *Ann. Math.*, **44**, 610–627 (1943).

[9] Vietoris, L., *Math. Ann.*, **97**, 545–572 (1928).

[10] Y is ω–LC_1 if there exists a covering σ of Y by open sets such that every closed path lying entirely in a set of σ is contractible to a point in Y.

[11] Y has barycentric refinements if, given any covering $\{U_\alpha\}$, a refinement $\{V_\beta\}$ exists with the property that, if $\bigcap V_{\beta\lambda} \neq 0$, then $\bigcup V_{\beta\lambda} \subset U_\alpha$ for some α.

[12] $Y^I(y)$ is the space of all paths in Y which start at y, with the compact open topology. Fox, R., *Bull. Am. Math. Soc.*, **51**, 429–432 (1945).

[13] A universal covering space is one with $\chi(Y) = 1$. Clearly $\pi_1(Y) = 1$ implies $\chi(Y) = 1$.

[14] Hurewicz, W., and Steenrod, N., PROC. NAT. ACAD. SCI., **27**, 60–64 (1941).

ON A GENERALIZATION OF A JACOBI EXPONENTIAL SUM ASSOCIATED WITH CYCLOTOMY

BY H. S. VANDIVER*

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF TEXAS

Communicated December 15, 1949

Let $\alpha = e^{2i\pi/(p-1)}$, $\zeta = e^{2i\pi/p}$, p prime then the resolvent used by Lagrange[1] in his theory of the solution of cyclotomic equations, dating from the year 1770,

$$\sum_{a=1}^{p-2} \alpha^{c \text{ ind } a} \zeta^{ak} = \tau(\alpha^c, \zeta^k); \tag{1}$$

(ind a) is defined by $g^{\text{ind } s} \equiv s \pmod{p}$, g being a primitive root of p, was the starting point for a series of investigations which are still being carried on.

Closely associated (cf. (13)) with (1) is[2]

$$\psi_1(\alpha) = \sum_{k=0}^{p-2} \alpha^{-bk + (a+b)\text{ind }(g^k+1)} \tag{2}$$

$$\alpha^{(a+b)\text{ ind }(0)} = 0,$$

with a, b and c any integers, which has probably been employed even oftener than (1) in algebra and number theory. In view of the remarkable properties of (1) and (2), a number of generalizations have been proposed for them.

Cauchy[2a] generalized (2) to take on a form equivalent to the sum (11) or (22) of the present paper if we assume $n = 1$ in each of them. These two latter forms for n general provide the main topic for our work here.

In the year 1920, H. H. Mitchell[3] considered, as an extension of (2)

$$\psi(\alpha) = \sum_{k=0}^{p^n-2} \alpha^{-bk + (a+b)\text{ ind }(g^k+1)}, \tag{3}$$

where ind s is defined by the equation $g^{\text{ind } s} = s$ in the finite field $F(p^n)$ of order p^n, p an odd prime and n an arbitrary integer > 0, s being any element of $F(p^n) \neq 0$, $\alpha^{\mu \text{ ind }(0)} = 0$ for μ any integer. An equivalent expression was used by Kummer[4] in the case where b and $(a + b)$ are divisible by the integer $(p^n - 1)/l$ and p belongs to the exponent l, where l is an odd prime.

Davenport and Hasse[5] in 1935 generalized (1) in the form

$$\tau(\alpha^\mu, \zeta^\nu) = \sum_a \chi^\mu(a) \zeta^{\nu \text{ tr }(a)}, \tag{4}$$

where now a ranges over all the elements of $F(p^n)$, $\mu \not\equiv 0 \pmod{m}$, $\nu \not\equiv 0 \pmod{p}$, $\chi(a)$ represents a character defined by the multiplicative cyclic group formed by the non-zero elements in $F(p^n)$, a ranging as stated, with

$\chi^{\mu}(0) = 0$, $\mu \neq 0$, and $\chi^{\mu}(0) = 1$, $\mu = 0$. The symbol $tr(a)$ means that if we write in $F(p^n)$, with the c's in $F(p)$

$$a = c_0 + c_1\omega + \ldots + c_{n-1}\omega^{n-1},$$

where ω is a root of a polynomial $f(x) = 0$ of degree n with coefficients in $F(p)$, also irreducible in $F(p)$ then

$$tr(a) = \sum_{i=1}^{n} (c_1 + c_2\omega^{(i)} + \ldots + c_{n-1}\omega^{(n-1)(i)}),$$

where $\omega^{(1)} = \omega, \omega^{(2)}, \ldots \omega^{(n)}$ are the n conjugate roots of $f(x) = 0$ in $F(p^n)$. The τ number of (4) will be termed a *Generalized Lagrange Resolvent*.

In the case where p belongs to the exponent n, modulo m, where α^m is an mth root of unity, a number equivalent to (4) was introduced and studied at length by Stickelberger[6] in 1890 in a celebrated paper.

In a recent article, Hua and Vandiver employed[7] (4) in examining the number of solutions of the equation

$$c_1x_1^{a_1} + c_2x_2^{a_2} + \ldots + c_sx_s^{a_s} + c_{s+1} = 0, \tag{5}$$

where the a's are integers such that $0 < a < p^n - 1$; $s \geqq 2$ for $c_{s+1} \not\equiv 0$ and $s > 2$ for $c_{s+1} = 0$ the c's being given elements of $F(p^n)$ and

$$c_1c_2 \ldots c_sx_1 \ldots x_s \not\equiv 0. \tag{5a}$$

If N is the number of solutions of (5) under the conditions just stated then they found

$$N = \frac{(p^n - 1)^s}{p^n} + \frac{T}{p^n}, \tag{6}$$

where

$$T = \sum_{\substack{a \bmod k \\ a \neq 0}} \prod_{i=1}^{s} \sum_{\chi_{k_i}} \chi_{k_i}(c_i^{-1}a^{-1})\tau(\chi_{k_i})\zeta^{-tr(ac_{s+1})}. \tag{6a}$$

Here

$$\tau(\chi_{k_i}) = \sum_{b \bmod K} \chi_{k_i}(b)\zeta^{tr(b)}$$

where $\chi_k(b)$ is a special k_ith character each element of which satisfies $x^k = 1$, $p^n - 1 \equiv 0 \pmod{k_i}/k_i = (a_i, p^n - 1)$, K stands for $F(p^n)$. In another recent paper, Weil[8] examined the number of solutions M of (5) *without the restriction, included in our* (5a), *that*

$$x_1x_2 \ldots x_s \not\equiv 0.$$

If $c_{s+1} = 0$ then in our notation he obtained if $r_i \not\equiv 0 \pmod{k_i}$,

$$M = p^{n(s-1)} + (p^n - 1)\sum_{\chi_{k_i}}^{s} \prod_{i=1} \chi_{k_i}(c_i^{-1})j(\alpha) \tag{7}$$

where if $\psi_{k_i}(b) = \Theta^{2i\pi r_i/k_i}$ then $\sum_i r_i/k_i \equiv 0 \pmod 1$,

$$j(\alpha) = \sum_{v_i} \chi_{k_1}(v_1)\ldots\chi_{k_s}(v_s) \tag{7a}$$

and where $v_1, v_2, \ldots v_s$ each range independently over all elements of $F(p^n)$ such that $1 + v_1 + \ldots + v_{s-1} = 0$, ((7a) was given in another form by Cauchy[2a] for $n = 1$ and by (11) of the present paper for n general).

He also obtained an analogous expression for M when $c_{s+1} \neq 0$.

If we include the case where $x_1x_2\ldots x_s = 0$ we may obtain a form analogous to our (6) by using[7] (the relation just above (11) of that paper with the restriction $x \neq 0$ omitted).

$$\sum_{x \in K} \zeta^{tr(cx^k)} = \sum_{y \in k}\sum_{\chi_k} \chi_k(c^{-1}y)\zeta^{tr(y)}$$

where we now define $\chi_k(0)$ as in connection with (4), noting that there we are expressing each of the $\chi_k(a)$, a fixed, as a power of a primitive one of them. This gives if N_1 is the number of solutions of (5) without the restriction

$$x_1x_2\ldots x_s \neq 0$$

$$N_1 = p^{n(s-1)} + \frac{T_1}{p^n} \tag{8}$$

where

$$T_1 = \sum_{a \in K} \prod_{i=1}^{s} \sum_{\chi_{k_i}} (c_i^{-1}a^{p^n-2})\tau(\chi_{k_i})\zeta^{tr(ac_{s+1})}$$

Our object in the present paper will be to develop some of the simplest properties of $j(\alpha)$ as given in (7a), without particular reference to the solution of equations in a finite field, except in connection with the proof of (15) below. We shall have occasion to employ these properties in subsequent papers.

A question of notation here assumes some importance. We may retain the notation of characters $\chi_k(w)$ in dealing with $j(\alpha)$ but since we are dealing with a *cyclic* group it seems more convenient to employ in place of $\chi_k(w)$ the explicit expression

$$\rho^{a \text{ ind } w} \tag{9}$$

where ρ is a primitive kth root of unity and (ind w) is defined as in (3). *If we adopt the latter, we immediately connect up our work to many results in the*

vast literature of cyclotomy. For example, Cauchy (who used I(a) for ind a), Kummer, Kronecker, Weber, Bachmann, Mertens and Landau, among others, used it. Also the fundamental property of characters

$$\chi(h_1)\chi(h_2) = \chi(h_1h_2)$$

is at all times in the foreground when we transform expressions of the type (9), since

$$\text{ind } h_1 + \text{ind } h_2 \equiv \text{ind } h_1h_2 (\text{mod } (p^n - 1)). \tag{10}$$

We shall employ (9) in the present paper. From what we have just explained it will be easy to translate any of our formulas into the notation of characters if so desired.

We then consider the exponential sum

$$\psi = \psi(\alpha_1^{\mu_1}, \ldots, \alpha_s^{\mu_s}) = \sum_{a_1, a_2, \ldots, a_s} \alpha_s^{\mu_s \text{ ind } A} \prod_{i=1}^{s-1} \alpha_i^{\mu_i \text{ ind } a_i} \tag{11}$$

where $\alpha_i = e^{2i\pi/m_i}$, $i = 1, 2, \ldots, s$; m_i is a divisor of $p^n - 1$; p an odd prime; μ_i, $i = 1, 2, \ldots s$, integers $\geqq 0$; ind a_i for $a_i \neq 0$ is defined by $g^{\text{ind } a_i} = a_i$ in $F(p^n)$, $i = 1, 2, \ldots s$,; g is a primitive root of $F(p^n)$; $\alpha_i^{\mu_i \text{ ind } (0)} = 0$; each a_i ranges independently over each element of $F(p^n)$, with

$$A = 1 - \sum_{i=1}^{s-1} a_i. \tag{11a}$$

The number ψ will be termed a *generalized Jacobi sum*. When (12a) holds m_i) it equals the $j(\alpha)$ of Weil. It was proved by Weil[8] (pp. 501–502) that

$$\psi\bar{\psi} = p^{n(s-2)} \tag{12}$$

provided

$$\left.\begin{aligned} &\alpha_i^{\mu_i} \neq 1, \; i = 1, 2, \ldots, s \\ &\prod_{i=1}^{s} \alpha_i^{\mu_i} = 1. \end{aligned}\right\} \tag{12a}$$

It was also shown by Weil[8] (p. 501; that for $\alpha_i^{\mu_i} \neq 1$; $i = 1, \ldots s$,

$$\psi = \frac{\prod_{i=1}^{s} \tau(\alpha_i^{\mu_i})}{p^n}, \tag{13}$$

provided $\alpha_1^{\mu_1}\alpha_2^{\mu_2}\ldots\alpha_s^{\mu_s} = 1$, where $\tau(\alpha_i^{\mu_i}) = \Sigma_a \alpha_i^{\mu_i \text{ ind } a}\zeta^{Tr(a)}$. Since for $a \neq 0$ in $F(p^n)$, there is an h such that $g^h = a$, we examine the following sum, which equals ψ or possibly its negative (cf. the derivation of our relation (24) which follows),

$$\psi_1 = \sum_{h_1, h_2, \ldots h_s} \alpha_s^{\mu_s \operatorname{ind} H} \prod_{i=1}^{s-1} \alpha_i^{\mu_i h_i}, \tag{14}$$

with

$$H = -1 - g^{h_1} - g^{h_2} \ldots - g^{h_{s-1}}.$$

We now proceed to write (14) in another more symmetrical form, namely

$$\psi_1 = \sum_{j_1, j_2, \ldots j_s} (j_1, j_2, \ldots, j_s) \prod_{i=1}^{s} \alpha_i^{\mu_i j_i}, \tag{15}$$

where $(j_1, j_2, \ldots, j_s)$ is the number of solutions of

$$1 + g^{j_1 + m_1 \gamma_1} + g^{j_2 + m_2 \gamma_2} \ldots + g^{j_s + m_s \gamma_s} = 0 \tag{16}$$

in the γ's, where γ_i ranges over the integers $0, 1, \ldots, m_i'$ with $m_i m_i' = p^n - 1$, and $i = 1, 2, \ldots, s$. To show that (14) and (15) are equal, we consider (14) and collect the terms in a certain way. Consider the residues of h_i modulo m_i and write for fixed j_i $h_i = j_i (\text{mod } m_i)$; $i = 1, 2, \ldots s-1$. Also assume that for a fixed j_s that ind $(-1 - g^{h_1} \ldots - g^{h_{s-1}}) \equiv j_s \pmod{(p^n - 1)}$. The number of terms of this character in (14) is exactly $(j_1, j_2, \ldots, j_s)$ hence we have (15). From this it is obvious that we have for a fixed set $j_1', j_2', \ldots j_s'$, using $\sum_{\mu_i} \alpha_i^{k\mu_i} = 0$ for $k \not\equiv 0 \pmod{m_i}$ and m_i otherwise

$$(j_1', j_2', \ldots j_s') \prod_{i=1}^{s} m_i = \sum_{\mu_1, \ldots \mu_s} \psi_1(\alpha_1^{\mu_1}, \alpha_2^{\mu_2} \ldots \alpha_s^{\mu_s}) \prod_{i=1}^{s} \alpha_i^{-\mu_i j_i'}, \tag{17}$$

where μ_i ranges over $0, 1, \ldots, (m_i - 1)$ with $i = 1, 2, \ldots s$.

We note that if

$$\tau^{p^n - 1} = (\sum \alpha^{\mu \operatorname{ind} a} \zeta^a)^{p^n - 1},$$

then this number is unaltered by the substitution (ζ / ζ^k) hence it contains α only. Also it is known that we have

$$\tau(\alpha^{p^t}) = \tau(\alpha). \tag{18}$$

Further we may show that

$$\psi(\alpha_1^{\mu_1 p^t}, \alpha_2^{\mu_2 p^t}, \ldots, \alpha_s^{\mu_s p^t}) = \psi(\alpha_1^{\mu_1}, \ldots, \alpha_s^{\mu_s}). \tag{19}$$

For we note that if

$$\beta_1, \beta_2 \ldots \beta_d, \text{ with } p^n - 1 = d, \tag{20}$$

are the elements $\not\equiv 0$ of a finite field then

$$\beta_1^{p^t}, \beta_2^{p^t} \ldots \beta_d^{p^t} \tag{21}$$

are equal in some order to the elements in (20) since if

$$\beta_i^{p^t} = \beta_j^{p^t};$$

$$(\beta_i^{p^t})^{p^{n-t}} = (\beta_j^{p^t})^{p^{n-t}}, \qquad \beta_i = \beta_j,$$

hence

$$\sum_{a_1, a_2, \ldots a_s} \alpha_1^{p^t \mu_1 \text{ ind } (A)} \prod_{i=1}^{s-1} \alpha_i^{p^t \mu i \text{ ind } a_i}$$

$$= \sum_{a_1, a_2, \ldots a_s} \alpha^{p^t \mu_1 \text{ ind } B} \prod_{i=1}^{s-1} \alpha_1^{\mu i \text{ ind}}$$

where $B = A^{p^t}$, $b_i = a_i^{p^t}$. Also from (11a)

$$A^{p^t} = \left(1 - \sum_{i=1}^{s-1} a_i\right)^{p^t}.$$

By the use of the multinomial theorem, we find that

$$\left(1 - \sum_{i=1}^{s-1} a_i\right)^p = \left(1 - \sum_{i=1}^{s-1} a_i^p\right) + p\gamma,$$

where γ is in $F(p^n)$ and therefore

$$\left(1 - \sum_{i=1}^{s-1} a_i\right)^p = 1 - \sum_{i=1}^{s-1} a_i^p,$$

since if $c_1, c_1 \ldots c_{p-1}$ are the elements of $F(p)$ then $c_p = c_0$ in $F(p)$ and therefore $F(p^n)$. By proceeding in this manner we have

$$\left(1 - \sum_{i=1}^{s-1} a_i\right)^{p^t} = 1 - \sum_{i=1}^{s-1} a_i^{p^t},$$

and in view of our remark concerning (20) and (21) our relation (19) follows. Considering the well-known ψ for $s = 2$, investigators have found it convenient, as we did in (14), to employ numbers closely related to ψ which have many of the properties of ψ for $s = 2$. Hence we consider numbers related to the general ψ, but appear to retain properties of ψ such as (12).

Let m be the L. C. M. of $m_1, m_2 \ldots m_s$. Then $m \leqq p^n - 1$. Then we may write, if $m = t_i m_i$; $i = 1, s, \ldots s$, with $\alpha = e^{2i\pi/m}$,

$$\psi = \sum_{h_1, h_2, \ldots h_s} \alpha^{t_1 \mu_1 \text{ ind } H} \prod_{i=1}^{s-1} \alpha^{\mu i t i h i}. \tag{22}$$

For many purposes this will turn out to be a more convenient form for ψ than either (11) or (14). Obviously also we may employ $p^n - 1$ in lieu of m, and obtain an analogous form. In connection with (22) we note also that in lieu of it we may employ the simpler form

$$\psi = \sum_{h_1, h_2, \ldots h_s} \alpha^{\mu_s' \text{ ind } H} \alpha^D, \tag{23}$$

where μ_i' and μ_s' are any integers $\geqq 0$, and where $\sum_{i=1}^{s-1} \alpha^{\mu_i' h_i} = D$. Clearly (23) includes the form (22).

We now note that in (11) we may obtain a number related to ψ by writing $(-a_k')$ in lieu of a_k in (11) letting a_k range over the same set that a_k did if k is in the set $1, 2, \ldots s$. For this gives

$$\psi_2 = \alpha_k^{\mu_k \text{ ind } (-1)} \sum_{a_1, a_2, \ldots a_s} \alpha_s^{\mu_s \text{ ind } A'} \text{ II } \alpha_i^{\mu_i \text{ ind } a_i}, \tag{24}$$

when we write a_k for a_k', noting that

$$\alpha^{\mu_k \text{ ind } (-a_k')} = \alpha^{\mu_k \text{ ind } (-1) \, - \mu_k \text{ ind } a_k}$$

and

$$A' = 1 + a_k - \sum_{\substack{i=1 \\ i \neq k}}^{s-1} a_i.$$

Now ψ_2 in (24) obviously has the same property (12). In the same way ψ_3, the number obtained by changing all the negative signs in A has the same property as well as the number ψ_1 of (14).

We note also that since p is an odd prime $g^{(p^n-1)/2} = -1$ in $F(p^n)$, so that in (24)

$$\alpha_k^{\mu_k \text{ ind } (-1)} = \alpha_k^{\mu_k (p^n-1)/2}.$$

We have also by definition of α_{k_i}

$$\alpha_k^{p^n-1} = 1, \; \alpha_k^{(p^n-1)/2} = \pm 1, \tag{25}$$

the sign depending on the value of m_k. In another paper[9] the writer pointed out that when α is an mth root of unity, m composite, then a relation between exponential sums found by the sole use of

$$\alpha^{m-1} + \alpha^{m-2} + \ldots + 1 = 0 \tag{26}$$

does not in general yield as much information as when we replace α by an indeterminate x in said relation and reduce the resulting polynomials modulo $(x^m - 1)/(x - 1)$. Hence we shall apply this idea to several equations given above, in order to obtain relations which are in effect more general. If we take ψ in the form given in (22) and multiply it by its conjugate imaginary, assuming the conditions (12a), we find by the use of (26) only, the relation (12), by proceeding in a manner equivalent to Weil's[8] derivation of it; now assume that we take ind H positive or zero in (22) and the μ's positive integers such that (12a) is satisfied then

$$\psi(x) = \sum_{h_1 \ldots h_s} x^{h_{s+1} \text{ ind } H \prod_{i=1}^{s-1} x^{a_i h_i}}$$

is a polynomial with rational integral coefficients, as is also $\psi(x^m - 1)$. Noting that $\psi(\alpha^m - 1) = \psi(\alpha^{-1})$, we then have

$$\psi(x)\psi(x^m - 1) = p^{(s-1)a} + C(x)\frac{x^m - 1}{x - 1} \tag{27}$$

where $C(x)$ is a polynomial with rational integral coefficients. Other relations such as (18) may be treated similarly.

* Most of the material in this paper was discussed in a mathematical seminar at the University of Texas and I am indebted to O. H. Faircloth, C. A. Nicol and Milo Weaver for corrections and suggestions.

[1] *Abhandlungen der Akademie zu Berlin* 1770, 1771. (Cf. our footnote 5 also.)

[2] Jacobi, C. G. J., *Gesammelte Werke* Bd. VI, 254, 274; VII, 393–400. Jacobi does not seem to have given ψ_1 in an explicit form involving α alone, he employed (13) for $s = 2$ to obtain properties of ψ_1.

[2a] Oeuvres, (1), III, 112.

[3] *Trans. Am. Math. Soc.*, **17**, 167 (1916).

[4] *J. für die Math.* (Crelle) 121–36 (1852).

[5] *J. für die Math.* (Crelle), **172**, 153 (1935). They termed this number a Generalized Gauss sum. It is a direct generalization of our (1). The reference given in our first footnote to the number defined in our relation (1) is given by P. Bachmann (*die Lehre von der Kreisteilung*, Tuebner Leipzig und Berlin, 1921, p. 76) and also H. Weber (*Algebra*, vol. 1, French edition, Gauthier Villars, Paris, 1898, p. 623) the latter remarking, concerning the numbers in our (1) "connues sous le nom de resolvantes de Lagrange." Further, Hilbert (*Gesammelte Abhandlungen* Bd. 1, 225) calls our (1) a "Lagrangesche Wurzelzahl" and the "cyclotomic periods" associated with (1) a "Lagrangesche Normalbasis." The term "Lagrange resolvent" is also used in *Report on the Theory of Algebraic Numbers*, National Research Council of the National Academy of Sciences, Washington, vol. 1, p. 50 (1923). In view of these circumstances we prefer the term "Generalized Lagrange Resolvent" for (4), at least until it appears that the references given by Bachmann and Weber are incorrect. So far we have been unable to check the original sources.

[6] *Math. Ann.*, **37**, 321 (1890).

[7] These PROCEEDINGS, **35**, 97 (1949), relations (10) and (11). The relation (8) of this paper is another result from the collaboration of Dr. Hua and myself.

[8] *Bull. Am. Math. Soc.*, **55**, 500 (1949).

[9] These PROCEEDINGS, **35**, 686–90 (1949).

ON DISTORTION AT THE BOUNDARY OF A CONFORMAL MAP

By J. L. Walsh

Department of Mathematics, Harvard University

Communicated December 19, 1949

The object of the present note is to indicate the usefulness in the study of distortion at the boundary, of Carathéodory's theory of the conformal mapping of variable regions. The emphasis here is on method, for the results are in part known.[1] Our main result is

THEOREM 1. *Let R_w be a Jordan region of the w-plane not containing $w = \infty$ but containing the line segment $0 \leqq w < 1$, and whose boundary possesses at the point $w = 1$ forward and backward tangents making equal angles $\alpha/2$ (>0) with the negative direction of the axis of reals. Let the function $w = f(z)$ map the region $R_z: x < 1$ of the $z(= x + iy)$-plane onto R_w in such a way that we have $f(0) = 0, f(1) = 1$. Let the sequence x_n, $0 < x_n < 1$, approach unity. Then for z on any closed bounded set in R_z we have uniformly*

$$\lim_{n \to \infty} \frac{f[(1 - x_n)z + x_n] - f(x_n)}{1 - f(x_n)} = 1 - (1 - z)^{\alpha/\pi}, \tag{1}$$

$$\lim_{n \to \infty} \frac{1 - f[(1 - x_n)z + x_n]}{[1 - f(x_n)](1 - z)^{\alpha/\pi}} = 1. \tag{2}$$

The tangent line is merely the limit of the secant line through $w = 1$ as the second intersection with the curve approaches $w = 1$.

We first establish Theorem 1 for the case that R_w is symmetric in the axis of reals; here the cut $0 \leqq w \leqq 1$ is the image of the cut $0 \leqq z \leqq 1$. The function $z' = (1 - x_n)z + x_n$ maps R_z onto itself, so the function

$$w = f_n(z) \equiv \frac{f[(1 - x_n)z + x_n] - f(x_n)}{1 - f(x_n)} \tag{3}$$

maps the region R_z in such manner that we have $f_n(0) = 0$, $f_n'(0) > 0$, $f_n(1) = 1$, onto the region R_w^n in the w-plane found from R_w by stretching from the point $w = 1$ in the ratio $[1 - f(x_n)]:1$. The kernel of the sequence of regions R_w^n is a region R_w^0, namely an infinite sector of angle α with vertex $w = 1$ and bisected by the axis of reals. The function $w = f_0(z)$ which maps R_z onto R_w^0 with $f_0(0) = 0, f_0'(0) > 0$, is the second member of (1), and equation (1) now follows by Carathéodory's theory of the mapping of variable regions. Equation (2) is an immediate consequence of (1).

If the points $\zeta_n = \zeta_n' + i\zeta_n''$ of R_z lie in an infinite sector S with vertex $z = 1$, of angle less than π and symmetric in the axis of reals, we set

$$x_n = \zeta_n', \qquad z_n' = i\zeta_n''/(1 - x_n), \tag{4}$$

so the points z_n' are collinear with $z = 1$ and ζ_n, and lie in S on the axis of imaginaries, hence lie on a closed bounded subset of R_z. Equation (2) with $z = z_n'$ and $\zeta_n \to 1$ thus implies

$$\lim_{n \to \infty} \frac{1 - f(\zeta_n)}{[1 - f(x_n)][1 - z_n']^{\alpha/\pi}} = 1. \tag{5}$$

In particular if the angle $arg(1 - \zeta_n)$ approaches a limit γ, so does $arg(1 - z_n')$, and from (5) we have

$$\lim_{n \to \infty} arg[1 - f(\zeta_n)] = \alpha\gamma/\pi. \tag{6}$$

Conversely, (6) implies $arg(1 - \zeta_n) \to \gamma$. These equations merely express the well-known fact (Lindelöf) that cuts with tangents at $z = 1$ and $w = 1$ are transformed into cuts with tangents at $w = 1$ and $z = 1$, and angles at those points are transformed proportionally.

We now differentiate (1) with respect to z and again make the substitution (4); we have

$$\lim_{n \to \infty} \frac{f'(\zeta_n)(1 - x_n)}{[1 - f(x_n)](1 - z_n')^{-1 + \alpha/\pi}} = \alpha/\pi. \tag{7}$$

Division of (7) by (5) member by member, with use of the equation $(1 - x_n)(1 - z_n') = 1 - \zeta_n$ now yields

$$\lim_{n \to \infty} \frac{f'(\zeta_n)(1 - \zeta_n)}{1 - f(\zeta_n)} = \alpha/\pi, \tag{8}$$

an important relation due to Visser. It is to be noted that if $arg(1 - \zeta_n)$ approaches the limit γ, then it follows from (6) and (8) that we have

$$\lim_{n \to \infty} arg[f'(\zeta_n)] = \gamma(\alpha - \pi)/\pi. \tag{9}$$

Equation (6) implies $arg(1 - \zeta_n) \to \gamma$ and hence (9).

Theorem 1 with various corollaries is now established for the case that R_w is symmetric in the axis of reals. The proof is not valid without that assumption, for we have essentially used the fact that $f_n'(0) = f'(x_n)(1 - x_n)/[1 - f(x_n)]$ is positive, or at least approaches a positive limit.

It is a consequence of (6) and (8) that the particular map $M_1: w = f(z)$ of R_w has *property A*, namely that angles at $z = 1$ or $w = 1$ bounded by curves in the given regions not tangent to the boundary but with continuously turning tangents are transformed proportionally, and the transformed angles are also bounded by curves with continuously turning tangents.[2] It follows that *any* map M_2 of R_w onto a half-plane so that $w = 1$ is invariant has property A, for the map $M_1(M_2^{-1})$ is a map of a half-plane onto a half-plane involving at $w = 1$ equality of angles bounded by curves not tangent to the boundary, hence by Schwarz's principle of symmetry this

map is conformal. Any map of R_w onto an infinite sector with vertex $w = 1$ and with $w = 1$ invariant can be accomplished by mapping R_w onto a half-plane with $w = 1$ invariant, followed by a transformation $w' = 1 - \mu(1 - w)^\beta$, $\beta > 0$, hence also possesses property A.

We have now proved, for a symmetric angle,

THEOREM 2. *On a Jordan curve C let M be an angle with vertex V at which C possesses forward and backward tangents. When the interior of C is mapped one-to-one and conformally onto the interior of a Jordan curve C' so that the sides of M in the neighborhood of V correspond to line segments, the map possesses property A, namely that any Jordan arc interior to C except for an end-point at V but with continuously turning tangent and not tangent to C is transformed into a Jordan arc interior to C' except for an end-point at the image of V but with continuously turning tangent and not tangent to C', and conversely. Angles at V are transformed proportionally.*

Let R_w in Theorem 1 no longer be symmetrical; we show that any map of R_w onto an infinite sector possesses property A. We assume, as we may do with no loss of generality, that R_w lies in an infinite sector with vertex $w = 1$ and of angle less than π. If R_w is bounded in part by a line segment terminating in $w = 1$, we may assume that no point of that line is in R_w, and we choose that segment on $-\infty < w \leqq 1$, reflect R_w in the axis of reals, and map the new region consisting of R_w plus its reflection onto a half-plane; this map transforms R_w into a quadrant and has property A. If R_w is not bounded in part by a line segment terminating in $w = 1$, we draw a line segment to the point $w = 1$ which is a cut for the exterior of R_w. A Jordan region R_w' bounded in part by this cut and in part by a side of the given angle contains R_w; any map of R_w' onto an infinite sector with vertex $w = 1$ and with $w = 1$ invariant possesses property A, and carries R_w into a region bounded in part by a line segment terminating in $w = 1$. Any further map of the latter region onto an infinite sector with vertex $w = 1$ and with $w = 1$ invariant possesses property A. Theorem 2 is completely established. Theorem 2 is due to Visser; the method of using angles bounded in part by straight lines is due in somewhat different form to Carathéodory.

We are now in a position to complete the proof of Theorem 1, where R_w is no longer symmetric. We choose an arbitrary sequence $x_n \to 1$, $0 < x_n < 1$, as before, and define the region R_w^n as the image of R_z under the transformation

$$w = f_n(z) \equiv \frac{f[(1 - x_n)z + x_n] - f(x_n)}{|1 - f(x_n)|} \cdot \frac{|f'(x_n)|}{f'(x_n)}, \qquad (10)$$

whence $f_n(0) = 0$, $f_n'(0) > 0$. The segment $0 \leqq z \leqq 1$ is a cut in R_z, and from Theorem 2 we have $\lim_{n \to \infty} f_n(1) = 1$. The kernel of the regions R_w^n is again R_w, so we conclude under the present hypothesis equations (1), (2),

and with the notation (4) conclude also (5), (6), (7), (8), (9). We have not merely established Theorem 1 but likewise

COROLLARY 1. *Under the conditions of Theorem 1, let S be a closed infinite sector less than π with vertex $z = 1$, S except for its vertex lying in R_z. If the sequence ζ_n lies in S and $\zeta_n \to 1$, then we have (8). In the notation (4) we have (5) and (7); if $\arg (1 - \zeta_n) \to \gamma$, we have (6) and (9).*

We remark that this proof makes use of well-known results on the topological character of a conformal map, especially the images of cuts in a region, but assumes no previous results on transformation of angles on the boundary. Even in our use of maps of variable regions we do not need the classical Verzerrungssatz, for the point $w = 1$ is accessible from the exterior of R_w, and for n sufficiently large all the regions R_w^n leave uncovered a Jordan region near $w = 1$ exterior to R_w; under a suitable linear transformation the set of functions $f_n(z)$ becomes a bounded set.

Results on the higher derivatives of $f(z)$ follow readily by differentiation of equations (1) and (2).

Our proof of Theorems 1 and 2 and Corollary 1 does not essentially depend on an assumption that R_w is a *Jordan* region. It is sufficient if the boundary of R_w possesses forward and backward tangents, in the sense that there exist two half-lines terminating in $w = 1$, and given two arbitrary closed sectors with vertex $w = 1$ containing those half-lines in their interiors, there exists a neighborhood of $w = 1$ in which all boundary points of R_w lie in those sectors; it is naturally assumed that this property is not possessed by a single half-line terminating in $w = 1$ and an arbitrary closed sector containing it. As another example, in which the boundary of R_w does not possess forward and backward tangents at $w = 1$, let R_w consist of the half-plane $u < 1$ (where $w = u + iv$) plus an infinite set of canals in $u \geqq 1$. These canals are to be non-overlapping, are to abut on the line $u = 1$ in segments whose mid-points are $v = v_n (\to 0)$, where the lengths of the segments are, respectively, $1/v_n^2$. Of course the canals need not be bounded by Jordan arcs, nor need R_w be symmetric in the axis of reals. It will be noted that the transformation $w' = 1 - (1 - w)^{1/2}$ transforms R_w onto a region of the w'-plane which leaves a quadrant with vertex $w' = 1$ uncovered, and the point $w' = 1$ is accessible from the exterior. Of course R_w may also be modified by the subtraction of suitable subregions of $u < 1$ adjacent to the boundary $u = 1$.

The purpose of the present note is to indicate a method (namely the use of Carathéodory's theory of variable regions) rather than to emphasize the extensive applications of the method, which are reserved for another occasion. For instance if R_w is a suitably chosen region with zero angle at $w = 1$, properly determined similarity transformations of R_w define regions R_w^n whose kernel is an infinite strip bounded by two parallel lines. Again, for certain regions R_w the kernel of a sequence of regions R_w^n defined by (3)

depends on the geometric nature of R_w and varies with the choice of the points x_n or $f(x_n)$; for instance R_w may consist of the interiors of infinitely many circles approaching the point $w = 1$ and joined by canals; information regarding the transformation of R_w can still be obtained by the present method.

[1] A recent summary of results on distortion at the boundary, with detailed references to the literature, is given by Gattegno, C., and Ostrowski, A., Mémorial des sciences mathématiques, fascicules 109 and 110 (1949).

[2] Of course property A implies that angles in the given region of either plane (assumed merely to exist as angles between tangents) are transformed proportionally into angles in the other plane.

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The Publication will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be announced, giving the relation of the paper to previous publications of the author or others and, in addition, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work, long tables of data should be avoided, but authors should be prepared in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though a few complicate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

REFERENCES to literature, numbered consecutively, will be placed at the end of the article and short footnotes should be avoided. It is suggested that references to periodicals be furnished in some detail and in general accordance with the standard adopted for the Subject Catalogue of the International Catalog of Scientific Literature, viz., name of author, with initials following (ordinarily omitting title of paper), abbreviated name of journal, with place of publication, series (if any), volume, inclusive pages, year. For example: Montgomery, T. H., *J. Morph. Boston*, 15, 781–815 (1911); or, Wheeler, W. M., *Koenigsberg, Schr. physik. Ges.*, 1–142 (1914).

MANUSCRIPTS may be sent to the Managing Editor, Edwin Bidwell Wilson, Harvard School of Public Health, 641 Huntington Ave., Boston 15, Mass., only by MEMBERS of the ACADEMY or by chairmen of the sections of the NATIONAL RESEARCH COUNCIL who will assume responsibility for the propriety and scientific standards of the paper. Articles are limited to six printed pages; excess space is billed to the author at $10.00 per page.

PROOF will ordinarily be sent,—with the understanding that charges for author's corrections shall be billed to him. Authors are therefore requested to make final revisions on the typewritten manuscripts.

REPRINTS should be ordered at the time of [illegible] manuscript. They will be furnished to authors at cost, approximately as follows:

SCHEDULE OF RATES FOR REPRINTS, CARRIAGE CHARGES, ETC.

No. Copies	50	100	150	200
4 Pp.	$3.40	$4.10	$4.61	$5.12
8 Pp.	7.12	8.21	9.24	10.26
Covers	5.76	6.50	7.24	7.98

CONTENTS.

Page

Determination of Water, Ash and Silicon Dioxide Volumetry in Organic Substances.
By [illegible] 67

Apparatus for Rapid Simple Analysis of Carbonaceous Gas Residues.
By Ernest Kopke 73

Determination of Iron Sulphides in Coal Residues.
By George Wilson and Paul E. Brown 86

Determination of Total Sulphur in Coal.
By George Wilson and [illegible] 89

Analysis of Sulphuric Acid and Nitrogen Oxides. VI. Absorption Data.
By Frederick Kershaw, Reginald Sibley and William J. Rabbitt 93

Comparative Quantitative Estimation of Reducing Substances in Dried Urine.
By Henry C. Sherman, Gustavus S. Fraps and Mary E. Bell 109

Determination of Ammonia in Coal Gas.
By T. Y. Thomas 119

Colorimetric Detection of Small Quantities of Nitrites.
By Leon P. Kinnicutt and Reginald K. Mitchell 124

Polysulphides.—Sulphur Estimation in Presence of Polysulphides.
By Edward Eggar and John Da Cruz 132

Determination of Ammonia, Nitrites, Nitrates and Chlorides in Water.
By Conrad C. Mueller 136

Manganese.—Volumetric Determination of Manganese and Chromium.
By Léopold Ratz 150

Phosphorus.—Rapid Estimation of Phosphorus in Iron, Steel, &c.
By Claude Chantby and [illegible] Schulze 157

Manganese.—Volumetric Determination of Chromium.
By J. Dupasquier 163

Manganese.—Use of a Platinum Anode in the Electrolytic Assay of Manganese Ores.
By R. A. Vesalius 164

Manganese.—On Manganese as the Reagent in Chemical Analysis.
By J. L. Walsh 168